ESSAYS ON K-12 EDUCATION IN DEVELOPING COUNTRIES - CAUSES, CONSEQUENCES AND IMPEDIMENTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Naveen Sunder

May 2019

ESSAYS ON K-12 EDUCATION IN DEVELOPING COUNTRIES - CAUSES, CONSEQUENCES AND IMPEDIMENTS

Naveen Sunder, Ph.D.

Cornell University 2019

This dissertation focusses on the causes and consequences of low human capital accumulation among children and adolescents in developing countries in Asia and Africa. My analysis is divided into three separate papers that explore different aspects of this research agenda. Broadly, chapter 1 of the dissertation describes the intergenerational learning impacts of a national–level school construction policy in India. Chapter 2 picks up where chapter one ends – using unique 17–year panel data from Senegal and Madagascar this analysis shows that children who perform better on learning tests in second grade have better later life (young adult) educational and learning outcomes (ages 23–25). In chapter 3, I take a step back and explore the role of one of the biggest impediments of school learning in the developing world – child marriage, which accounts for a large proportion of school dropout among girls in developing countries.

In the first chapter of my dissertation I examine the impact of a national-level school construction program in India on educational outcomes of direct beneficiaries and their children. Between the years 1993–2004, the District Primary Education Programme (DPEP) served over 50 million children and prioritized districts with below–average female literacy rates. I use a fuzzy regression discontinuity design to estimate the causal impact of the programme by comparing

outcomes of school–age children in districts on either side of the average female literacy cutoff. To uncover the difference in timing of programme implementation across districts, I use unique archival information that I collected and digitized. The results show that DPEP increased school access, enrollment, literacy and years of education for both male and female direct beneficiaries. I then provide one of the first evidence of intergenerational effects of a school construction policy. Using test score data spanning the years 2007–2014, I find that children whose mothers were DPEP beneficiaries had higher scores on math (0.18 S.D.), vernacular (0.19 S.D.) and English (0.09 S.D.) tests. Daughters test scores went up by more than 10 to 15 percentage points higher than that of sons. Fathers DPEP exposure had no effect on childrens learning. I find evidence that the intergenerational impacts may be mediated through mothers increased bargaining power, higher investments in childrens education and better health/health related behaviors.

In the second chapter, I (along with my co–authors) study the determinants of human capital outcomes of young adults in Madagascar and Senegal, employing a production function approach. Using unique and comparable long–term panel data sets, which span more than 15 years, from both countries, we find that test scores in second grade are strong predictors of school attainment and French/math skills of individuals in their early twenties. The association between second–grade skills and later–life outcomes is stronger among girls than boys, and likewise, stronger for math than French test scores. These findings highlight the importance of not falling behind during early school years, as it can lead to worse long–term outcomes, particularly for vulnerable groups like girls. We also find that height, a proxy measure of childhood health and nutritional status, does not affect the magnitude and significance of the early childhood test score variable, and also has an independent effect on the test scores of young adults in Senegal.

Chapter 3 analyzes whether Ugandan women who marry at younger ages fare differently on a wide range of later life outcomes than women who marry at later ages. Using a nationally representative dataset, I identify the causal impacts of womens marriage age by using their age at menarche, a plausibly exogenous variable, as an instrumental variable. Results indicate that a one year delay in marriage leads to higher educational attainment (0.5–0.75 years), literacy (10 p.p.) and labour force participation (8 p.p.) among women. I also explore intergenerational effects of later marriage and find that the children of mothers who marry later have higher BMI (0.11 kg/m2) and hemoglobin levels (0.18 g/dl), and they are also less likely to be anemic (4 p.p.). Finally, I present evidence that suggests that the observed effects might be mediated through an enhancement womens agency within their household and positive assortative matching in the marriage market. By pointing to the beneficial consequences of delaying marriage, this research calls for concerted policy action to prevent child marriage.

BIOGRAPHICAL SKETCH

Naveen Sunder was born in Chennai (India) in 1988. He did his Bachelors and Masters in Economics in the University of Delhi, after which he worked for three years as a Research Assistant in the Indian School of Business (ISB), and the Jameel Poverty Action Lab (J–PAL). He likes soccer (football!), cricket, TV shows, movies and lazing around in his room.

I dedicate my thesis to my grandparents – Krishnaswami, Jayalakshmi, Ramadoss & Janaki. They have, and will always remain, a huge source of inspiration for me.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to David Sahn, the chairperson of my PhD dissertation committee, Prabhu Pingali and George Jakubson. They have provided invaluable advice, and without their support this PhD would not have been possible. I would like to thank the staff of the Cornell Economics department for providing timely and prompt institutional support throughout my time here. Special thanks goes to the Tata Cornell Institute (TCI) for providing me with a platform to develop my professional and academic skills in addition to introducing me to a great community of inter–disciplinary researchers that have broadened my horizons. TCI has also doubled up as home away from home – I have made many friends that I will cherish all my life.

I would also like to acknowledge the role of Mathew Abraham, Anaka Aiyar, Jessica Ames, Vidya Bharathi, Soumya Gupta, Hautahi Kingi, Yanan Liu, Sreemati Mukherjee, Tanvi Rao, Caroline Schimanski, Smriti Sharma, Gargi Wable and Vidya Vemireddy – each of them has provided me with support, encouragement and advice that has helped me get through challenging phases in my PhD. In particular, none of this would have been possible without the timely help that David Coriell provided in 2014 – I will always be indebted to him.

Special thanks goes to Prateek Bansal, who has been a friend, mentor and confidante – we met by accident (a sublet listing on craigslist!) but have now become best of friends. I have learnt a lot from him, most importantly the critical role of discipline in shaping a successful research career. I will for always cherish the candid conversations we have had on a variety of topics, including life, sports and academics. Thanks for always being there for me!

Averi Chakrabarti is the bedrock on which my PhD stands – she has accompanied me through this journey and has been one of my closest ally both professionally and personally. She has always spoken her mind and given me the advice that I need (not always the one that I want to hear!). Her constant encouragement and (tough) love kept me focussed and motivated. Thank you for putting up with my quirks and ranting, and for being by my side through thick and thin!

Last but surely not the least, I want to express my gratitude to my extended family, grandparents, uncle (Seshadri) and my parents. They have all had an immense role in shaping my personality. I am thankful to them for always stressing on the importance of helping others and for being an upstanding citizen above all else. In particular, I want to thank my parents, who through their hardwork and sacrifices over the years have made all of this possible – I would not be here without their efforts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

PARENTS' SCHOOLING & INTERGENERATIONAL HUMAN CAPITAL: EVIDENCE FROM INDIA

1.1 Introduction

Over the past few decades, there have been marked improvements in schooling availability across most parts of the world (Barro and Lee, 2013). Even though this has meant that more children attend school, it has not always been accompanied by improvements in learning levels in schools, especially in developing countries. For example, although India, the context of this study, has a net enrolment rate of close to 100 percent, only 43 (33) percent of sixth (seventh) grade Indian students could read a vernacular text at the second grade level, and only around a quarter of fifth grade students were able to solve a math (division) problem (ASER, 2016). Various studies have demonstrated the adverse effects of low childhood learning (Glewwe, 1996, Behrman et al., 2008, Kaila et al., 2019). This has lead to an enhanced focus on policies aimed at improving learning outcomes in schools. Although learning among children is a function of school, household and individual level inputs (Glewwe and Muralidharan, 2016), a bulk of the interventions addressing learning deficits have targetted school inputs. The findings of these evaluations are mixed – test scores have been found to be unresponsive to a number of these interventions (summarized in Kremer et al., 2013 and Muralidharan, 2013).

In this paper, I focus on the household inputs channel[1], and the intergenera-

[1] Studies have looked at different types of interventions in this respect - information provision (Jensen, 2010, Loyalka et al., 2013, Wang et al., 2014), conditional cash transfers (Behrman et al.,

tional effect that parents have in shaping children's educational outcomes (Black et al., 2005, Holmlund et al., 2011, Carneiro et al., 2013). In particular, I use data from India to evaluate if enhanced schooling access for parents, when they were of school–going age, not only improves their own educational outcomes, but if it also leads to improvements in their children's learning outcomes.

I do so by examining the direct and intergenerational impact of a national level school construction policy in India – District Primary Education Programme (DPEP). DPEP was implemented in 271 districts (in 18 states) between the years 1993 and 2004. This programme expanded school access by constructing primary and upper–primary schools, and was targetted towards districts with female literacy below the national average at the time (39.2 percent). This assignment mechanism creates a discontinuity in the probability of receiving the programme around the threshold of 39.2 percent female literacy. That is, the probability of receiving the treatment is much higher in districts just below this cutoff, as compared to districts just above this threshold. In implementing the Fuzzy RD design I use data–driven tools which estimate the causal impact within an optimal neighborhood around the RD cutoff. I employ two different approaches to constructing these neighborhoods, namely Mean Squared Error (MSE) approach and the Coverage Error Rate (CER) approach (Calonico et al., 2014, 2016) – the results are consistent across both approaches.

One of the innovative aspects of this paper is the use of unique archival data that I collected, which enables me to uncover differences in timing of programme implementation across districts, something that other analyses examining this programme have not done (like Azam and Saing, 2017, Khanna,

2009, Baird et al., 2011, Barrera-Osorio et al., 2011), scholarship programmes (Blimpo, 2014, Li et al., 2014) and other in–kind transfers (Oster and Thornton, 2011, Muralidharan and Prakash, 2017).

2015). These archival documents consist of programme expenditure information, field reports on implementation and other state/federal government documents monitoring DPEP progress. I triangulate information from these documents to uncover when a programme actually took effect in a treatment district. This enables me to accurately infer the start year of the programme in each of the 271 treatment districts, which I use in my empirical strategy[2] [3].

By comparing districts on either side of the RD cutoff, I establish that DPEP regions had higher rates of school construction[4]. This implies that during the programme years children of school–going age in DPEP districts experienced increased schooling infrastructure, as compared to children of the same age group in non–DPEP districts. Since the bulk of the schools constructed under the programme were primary and upper–primary schools (grades 1 to 7), children between the ages of 5 and 14 years are expected to benefit from DPEP. Since DPEP ended in 2004, only children born before 1999 could potentially benefit directly from the programme. The earliest cohort impacted by DPEP would vary from one district to another, depending on when DPEP took effect in the district. For example – If DPEP was implemented in a district in 1995, then people born between 1981 and 1999 would be the direct beneficiaries in this district. I find that both male and female direct beneficiaries of DPEP had higher enrolment, literacy and years of education, as compared to non-beneficiaries. This is in line with findings from other comparable studies (Duflo, 2001, Burde and Linden,

[2]I use 2004 as the uniform end year of the programme across the country. This is because the funding for the programme was cut in a phased manner between 2001 and 2004, implying the programme ended around late 2004.

[3]Since the non–DPEP districts did not receive the programme, there is no obvious start year of the programme in these districts. Therefore, I use the within–state average start year of all treatment districts as the start year of the programme in non–DPEP districts. This is needed to define a comparable control group – discussed in detail later

[4]However, there were no changes in the quality of schools in the DPEP districts. I measure quality using indicators on physical infrastructure, teacher quality, grants/incentives and school oversight.

2013, Kazianga et al., 2013).

In addition, I use school learning data from the Annual Survey of Education Report (ASER) data to provide one of the first estimates of the intergenerational impacts of a school construction policy. To do so, I use data from the years 2007 to 2014 and focus on a sample of children whose parents were of school going age during the DPEP years, but the children themselves did not directly benefit from the programme. To ensure this, I restrict my sample to children who started school after the programme had ended, that is, children who started school in/after the year 2005. Therefore, the intergenerational sample consists of children who satisfied two conditions – they started school in/after 2005 and they had at least one parent who was between 5 to 14 years of age during the DPEP years in their district.

Using an analogous RD framework (as above), I find that the children whose mothers were the sole beneficiaries of DPEP performed better on vernacular (0.19 S.D.), math (0.18 S.D.) and English (0.09 S.D.) tests, and were more likely to be enrolled in school and achieve smooth progression through grades, as compared to similar children born to comparable women in non–DPEP districts. In contrast to these results, I find no statistically significant positive impacts among the sample of children whose fathers (and not mothers) were exposed to DPEP. Thus, while both genders gained from direct exposure to the school construction programme, the results suggest that only women were able to transmit their benefits to the next generation. Like the evidence from several other studies (Desai and Alva, 1998, Persico et al., 2004, Case et al., 2005, Güneş, 2015), this result demonstrates the critical role played by mothers in their children's human capital development.

I conduct a number of falsification and robustness checks. In a falsification check, I show that people who were too old to directly benefit from the DPEP did not experience any of its gains. In line with expectations, in another falsification check I find that children of these people also do not show any DPEP impacts. Both approaches to RD estimation employed here require the researcher to specify the kernel function and polynomial form to be used in the estimation. I show that the results are robust to changes in the RD approach (MSE & CER), kernel function (triangular and epanechnikov) and polynomial form (linear and quadratic). In another robustness check, I use a stricter cutoff to define the potential direct beneficiary sample – I use the 5–12 years age group to define the direct beneficiary sample, rather than 5–14 years that I use in the main estimations[5]. Results remain robust to this change. In the main results, I use approaches to RD estimation that use data within a neighborhood around the RD threshold to estimate DPEP impacts. As a robustness check, I instead use the 2SLS IV technique, which imposes parametric assumptions on the whole data to estimate the impact coefficient. Although the value of the coefficients differ from the main results, the effects retain their sign and significance.

Finally, while the intergenerational effects that I observe can be plausibly attributed to parental educational attainment through DPEP, I conduct additional investigations to understand how individuals were able to use their additional education to impact their children's learning. I find that the women who benefitted directly from DPEP were healthier, had enhanced bargaining power and were investing more in their children's education as compared to similar women in non–DPEP districts.

[5] A lower cutoff of 12 years might be especially valid for girls since they tend to drop out of schools at younger ages than boys due to a variety of reasons – including onset of menarche and child marriage.

I contribute to this literature in the following ways. First, I add to the evidence on the important role played by parents, especially mothers, in shaping learning outcomes of children. This outlines the critical role that household inputs can play in complementing school factors to engender better educational outcomes. Second, this is one of the first papers to look at the intergenerational impact of a large scale school construction policy. Most papers have restricted their focus to the impact of such policies on direct beneficiaries (Duflo, 2001, Azam and Saing, 2017). Third, it evaluates the learning impacts of a school construction policy, something that is uncommon in the literature (Burde and Linden, 2013). Fourth, large parts of Central/Western Africa and South Asia have comparable educational indicators to what India had at the time of DPEP implementation. Therefore, results from this analysis could be informative for policy formulation in these contexts.

1.2 Literature & Background

This paper adds to the limited literature that estimates the effects of increasing the supply of schooling infrastructure on educational outcomes in developing countries by estimating intergenerational effects of such programmes. The bulk of this literature estimates the impacts on direct beneficiaries (Duflo, 2001, 2004, Handa, 2002, Alderman et al., 2003, Burde and Linden, 2013, Kazianga et al., 2013, De Hoop and Rosati, 2014, Giri and Shrestha, 2017).

Given that the focus of this analysis is on people who directly benefit from enhanced schooling opportunities, and their ability to impact their children's educational outcomes in the future, my study is closely related to the body of literature investigating the intergenerational links in education outcomes between parents and children. A bulk of this literature has found that parents' enhanced education has a direct positive impact on the educational (years of education and enrolment) and health outcomes of their children (Black et al., 2005, Sacerdote, 2007, Chou et al., 2010, Holmlund et al., 2011)[6]. In addition, it is plausible that better educational outcomes for the parent generation might lead to improved health for themselves (Amin et al., 2013, Agüero and Bharadwaj, 2014, Grossman, 2015, Grépin and Bharadwaj, 2015), higher age of marriage and age at first birth (Glick et al., 2015, Grant, 2015, Marchetta and Sahn, 2016), increased contraception usage and better antenatal care practices (Andalón et al., 2014, Johnston et al., 2015, Behrman, 2015), higher bargaining power for women (Lundberg and Pollak, 1993, Duflo, 2012, Samarakoon and Parinduri, 2015) and higher investment in children's education (Yoong, 2012), among other outcomes. These in turn might have a positive impact on their children's outcomes (intergenerational effect). Among these, in the current setting I find evidence for a positive impact on women's bargaining power, investment in children's education and own health/health related behaviors.

There so exist some papers that have evaluated the impacts of the DPEP policy, but only on socioeconomic outcomes for direct beneficiaries. Khanna, 2015 estimates the effect of this school expansion on the rate of return to education, while Azam and Saing, 2017 find that DPEP beneficiaries had higher enrollment, number of years of education and probability of completing primary

[6]This direct link between the educational outcomes of parents and their children has been studied extensively using many different empirical strategies, which include comparing twins and their children (Behrman and Rosenzweig, 2002, Bingley et al., 2005, 2009, Holmlund et al., 2011), natural experiments related to compulsory schooling laws/tuition fees/location (Black et al., 2005, Oreopoulos et al., 2006, Carneiro et al., 2013, Chevalier et al., 2013) and comparing outcomes between biological children and adopted children of the same parents (Sacerdote, 2002, 2007, Silles, 2017).

education. Using data from 42 districts that received the programme in phase one, Jalan and Glinskaya, 2013 find small effects on enrolment, that too mostly for socially disadvantaged groups. While there are some parallels between the current analysis and these previous DPEP papers, there are critical differences. I use detailed archive data to uncover the exact timing of the programme in the 271 treatment districts. This enables me to define the beneficiary sample more accurately and distinguishes this analysis from the aforementioned papers – which do not account for the different start dates of the program in this manner. Also, my results not only confirm that the program had positive impacts on the educational attainment of direct beneficiaries, but also establish that it had intergenerational impacts – which is another key innovation of my analysis.

1.2.1 DPEP Programme

The District Primary Education Programme (DPEP) was the flagship education programme of the Indian government in the 1990s. Implemented in a phased manner across the country, the programme was oriented towards achieving universal education through an increase in schooling infrastructure. Program rules stipulated that DPEP school construction would be targetted towards districts which had District Female Literacy Rates (DFLR) below the national average. The programme cutoff was chosen to be the national average female literacy at the time, which was 39.2 percent – based on the most recently available census data at the time (1991 Census). This programme allocation rule was largely followed. The central government also specified that DPEP would only be introduced in districts that had successfully implemented the Total Literacy Campaign (TLC), a programme that aimed at improving literacy levels across the country (Rao, 1993). Since the TLC had been implemented successfully across all districts in India by 1994, this criterion turned out to not matter for selection into the program (Jalan and Glinskaya, 2013). The program was introduced in 42 districts in 1993, and eventually extended to 271 districts across 18 states.

DPEP was financed by different levels of governments. The central government bore 85 percent of the costs with the support of donors like Official Development Assistance (ODA), the Royal Government of the Netherlands, the U.K. Department for International Development (DFID) and the World Bank. The remaining 15 percent was contributed by the state governments. To ensure that the new programme did not crowd out state funds that were already being spent on existing educational policies, the central government stipulated that states had to continue with their pre–existing non–DPEP expenditures. Since the DPEP funds were committed over and above the existing education budget, the programme represented a massive surge in education expenditure across the country.

1.3 Data

Below, I describe the different data sources that I use in this analysis.

1.3.1 Annual Survey of Education Report (ASER)

Pratham, an Indian non–profit organization, conducts national annual surveys to measure schooling and learning outcomes across rural India called the Annual Surveys of Education Report (ASER). These provide repeated cross sectional data on the educational profile of children aged five to 16 across the entire country. In my analysis, I use eight rounds of ASER data, spanning 21 states (see figure 1.2) for the years 2007–2014. ASER data contain student test scores for math, vernacular and English. An ASER test includes four questions, with each being more difficult than the previous question. The score assigned to a child indicates the level of difficulty that a child was able to solve/master. For example, on the reading section, a child gets a score of zero if she could not read anything; one if she is able to recognize alphabets, and two, three or four depending on whether she is at best able to read words, sentences or a paragraph respectively. I use this information to create a score variable which ranges from zero (if a child failed to answer any question correctly) to four (if she demonstrated the highest level of proficiency).

ASER also contains other information on additional educational outcomes for each child that is surveyed. Based on the available information, I create a grade–for–age variable that measures whether a child was held back at school or joined school at a later age than he/she should have. Here I use the fact that school starting age in India is typically 5–6 years. This variable takes a value of one if the child is on track in school, that is when age minus grade is at most six, and a value of zero otherwise (akin to Shah and Steinberg, 2017). For example – if a nine year old child is in grade three or four, she get a value of one. But if she is in grade 2, then she gets a value of zero. I create an additional indicator variables for whether a child has ever been enrolled in school.

One of the key advantages of using the ASER dataset is its national coverage. The sample size of each survey is large and it encompasses all rural districts in India[7]. Another unique aspect of the ASER data is that it measures educational

[7] See this link for further details on the ASER sampling strategy. In fact, this model has

achievement at home instead of schools. As a result, the sample includes children who have dropped out of school and children who have never enrolled, along with those that are currently enrolled in school. Additionally, the format of the tests, and the way they are administered and scored have remained uniform across different years and regions, thus facilitating spatial and temporal comparisons.

1.3.2 District Information System for Education (DISE)

For school characteristics, I use the District System for Education (DISE), a database of government schools across the country which contains information on the physical infrastructure and amenities present in each government school, as well as information on teachers, enrolment and other school–level covariates. This source compiles data provided by school headmasters on a yearly basis. Information provided by schools is verified at the cluster level and subsequently transferred to the district level. At this stage, the data is verified again before being aggregated, digitized and published. I use the DISE data for the year 2005, which is the earliest year of data that is publicly available.

I use this dataset in several different ways. The DISE data allows me to test whether the DPEP led to higher school construction in treatment regions during the years the program was in operation. The DISE data also enables me to examine treatment–control differences in school quality before, during and after the implementation of DPEP. Finally, using the DISE data, I cross–verify the accuracy of information I gather from government archival documents (described

been adopted in some African countries to measure learning levels among children – UWEZO Surveys below).

1.3.3 Archived Government Records

In order to isolate the timing of DPEP program initiation in different parts of the country, I take advantage of archived government records on the implementation of the program. Since the DPEP started in the early to mid 1990's, a large amount of the documentation pertaining to the programme was initially not in a digital format. Although, some of the documents have been recently digitized, a large amount of information still exists solely in paper format in libraries and other institutions. I gather data from these digital and paper files to infer information for programme districts across various states (a total of 271 districts). I discuss this in further detail in the next section.

1.3.4 District Level Health Survey (DLHS)

I obtain data on the educational attainment of DPEP's direct beneficiaries from the DLHS, a household–level survey conducted by the government of India. This survey collates statistics on a wide range of indicators related to household demographic characteristics, maternal and child health, and family planning. I draw upon two rounds of this data – rounds in 2007–2008 (wave 3) and 2012–2013 (wave 4). In addition, I also use this data to investigate the mechanisms that could potentially be responsible for DPEP's intergenerational effects.

1.3.5 Indian Human Development Survey (IHDS)

The IHDS is the other dataset that I use to probe the mechanisms that might have mediated the intergenerational impacts of DPEP. The IHDS is a nationally representative panel survey conducted by the University of Maryland in collaboration with the National Council of Applied Economic Research, New Delhi. The first round, IHDS–I, was conducted between November 2004 and October 2005 and covered 41,554 households across 33 states and union territories of India (Desai and Vanneman, 2005). In this analysis, I primarily use the female module that was administered to ever–married women between the ages of 15 and 49.

1.4 Empirical Strategy

Before discussing the details of the empirical strategy, I describe some important characteristics of the way I create the sample of direct and indirect DPEP beneficiaries (children of direct beneficiaries). It is to be noted that in some places the direct beneficiaries are referred to as the *parent* generation and the indirect beneficiaries as the *child* generation or intergenerational beneficiaries.

1.4.1 Sample Definition

In this section, I first describe the sample I use for my analysis. The direct beneficiaries of DPEP are those who were directly exposed to the program since they were of school–going age at the time of programme implementation. But, the focus of the analysis is on estimating the intergenerational effects of DPEP. Hence, my main sample constitutes the children of these direct beneficiaries.

Timing of the programme – DPEP Districts

DPEP was announced in 1993–94. Therefore, one way is to define everyone who was of primary school–going age at this point in time to be a direct beneficiary of DPEP, an approach used by Khanna, 2015. However, it is worth recognizing that DPEP constructed over 100,000 schools spread across 271 districts. Given the inherent challenges and the wide geographical spread of DPEP, the program began at different times in districts across the country. Using archived government documents[8], I find the exact year of initiation of DPEP in each treatment district, and use it to identify the direct beneficiary sample – people who were of primary school–going age at the time that DPEP was implemented in their district.

Given that I manually infer these start dates from these archival documents, I conduct a check to verify if this data holds up to further scrutiny. Using data from DISE, I plot the annual rate of growth of schools over time in select districts (figure 1.3). DPEP was meant to boost school construction in treatment districts. Therefore, one would expect a deviation from the long term pattern in the rate of growth of schools in or around the time that the programme was ac-

[8]I manually obtained the start date of the programme by triangulating information on programme expenditures, field reports on programme implementation and progress reports created at the state/central level to infer the time when school construction under DPEP actually began in the treatment districts. For example – I consider DPEP to have begun in a district if money has been received by the district(central documents), it has been spent by the state/district authorities (expenditure reports) and construction of schools has happened (progress reports). Therefore, when multiple pieces of information about DPEP implementation in a district provides a coherent narrative, I use this to infer the time when the programme was implemented there.

tually implemented in districts. On the graph (figure 1.3), the red line (dashed) represents the year when the DPEP programme was meant to begin in all the treatment districts (1993–94) and the black line (solid) shows the year in which the (aforementioned) government records suggest that the programme was implemented in. As the graphs illustrate, the actual spike in the rate of growth of schools in the districts is closer to the start year that I identify from the government documents (solid black line), rather than the uniform start year of 1993–94 (dashed red line). While the data for the underlying graph and the date from government archives (the black line) are from different sources, both pieces of information indicate that DPEP implementation began in different districts well after the central announcement of the program in 1993–94. For my analysis, I thus use the date inferred from the archives for each district (black line) to identify which individuals were exposed to the implementation of DPEP in that district.

Sample Selection – Non-DPEP Districts

I seek to identify the causal impacts of DPEP by comparing the prospects of individuals who were of school–going age in DPEP districts for the duration of the program, with the outcomes of comparable individuals in non–DPEP districts. If DPEP were assigned to these districts, some people by virtue of their age would have benefitted from it, while others would have missed out. The former is the comparison group in my analysis. Since the non–DPEP districts did not receive the programme, there is no obvious *start year* of the programme in these districts. Thus I need to understand when the program would have started in these districts should they have received it.

I use two different methods to assign a likely DPEP start year for the control districts. Under the first method, for each state I take the average starting year of DPEP in treatment districts within that state, and consider this year to be the start date for all control districts within the state. I use this in my main analysis. Another way to impute the starting year would be to assign all control districts the nationwide average start date among the treatment districts (instead of using the individual state averages). This method ignores the state (or regional) differences in implementation patterns across different districts. To show the robustness of my results, I replicate the main results using this alternative (national) definition of identifying the control group.

Direct Beneficiary Sample

DPEP was mostly geared towards the construction of primary and upper–primary schools (up to grade 7). Thus, the treatment and comparison sampels consist of individuals who were of primary school–going age in the treatment and control districts at the time of DPEP implementation. Primary school children tend to be between the ages of 5 and 10 years, but studies from India indicate that even children up to the age of 13 years might remain in primary school, mostly due to delayed school entry and/or uneven grade progression (Azam and Saing, 2017). As a result, I define the main sample of direct DPEP beneficiaries to include children who were between 5 and 14 years of age during the DPEP implementation years. I check the robustness of my results to an alternative definition that consider 5 to 12 years to be the relevant group.

Intergenerational Sample (Children of Direct Beneficiaries)

My main focus in this analysis is to identify the intergenerational effects of DPEP. In order to do this, I need to identify the children of DPEP direct beneficiaries (identified above) who themselves were not directly impacted by the programme. Given that the DPEP programme ended around late 2004[9], I only consider those who started school after this time period – specifically, I restrict the sample of children in my analysis to those who would have began their schooling in 2005 or later.

Identifying district of schooling

Ideally, an individual would likely be considered a DPEP direct beneficiary if he/she were of school–going age when they resided in a treatmetn district. The data I use to identify the impacts of DPEP on the direct beneficiaries, were collected in the year 2007 and later. While they include information on past educational attainment of direct beneficiaries, the surveys did not ask individuals to report their district of residence during their school–going years. I thus use individuals' current district of residence to assign treatment/control status to each individual (direct beneficiary) in my sample. The potential issue with this assignment mechanism is that, due to migration, the current district of residence may not be the same as the one that the individual lived in when they went to school.

One of the major sources of migration in India is post-marriage movement

[9]A different national program called Sarva Shiksha Abhiyan (SSA) was introduced in India around 2001–2002. While this program also aimed at expanding educational opportunities across the country, it differed from DPEP in that it wasn't targeted based on an allocation rule.

of women. In India, which is largely a patriarchal country, it is common for women to move to live with her husband's family after marriage which leads to systematic inter–district migration particularly for women. As a result, the current location of women may not always be an appropriate proxy for their past district of residence. There is however evidence that indicates that the majority of marriage–related migration occurs within and not across districts. Bloch et al., 2004 show that on average, a woman moves 21 miles after marriage. In 2001, the average size of an Indian district was close to 2,100 squared miles and thus it is highly likely that most post–marriage migration occurred within districts. Evidence from multiple nationally representative data sources point to this conclusion. Using National Sample Survey (NSS) data from 1983, 1987 and 1999, Topalova, 2007 finds that although nearly 60 percent of rural women report a change in their location after marriage, a very small proportion (7–8 percent) move across district boundaries[10]. In light of these statistics, I argue that it is reasonable to consider the current district of residence to be the district in which individuals went to school. Hence assigning treatment status based on the current district of residence is unlikely to lead to substantial misclassification errors.

Changes in District Boundaries

India has witnessed substantial administrative decentralization over the past two or three decades – the number of districts in the country has increased from 466 (in 1991) to 640 (in 2011). Given that the DPEP programme was implemented at the district level, it is crucial that I be able to link current district

[10]Using more recent data, Kone et al., 2017 show that overall inter district migration among women in India stands at about 9–10 percent. Almost 70 percent of this migration is due to marriage, but most of it (more than 3/4th) occurs within the same district.

definitions to districts boundaries that existed during the program years (1993–2004)[11]. This is important because the parent cohort went to school in old districts (1993–2004), and hence, whether or not they were received the benefits of DPEP would depend on their district of residence during program implementation.

Based on the discussion of historical changes in district boundaries in Kumar and Somanathan, 2009, I map districts in 2001 to their parent districts in 1991. While Kumar and Somanathan, 2009 only cover district changes that took place until 2001, I extend their analysis to similarly match districts that were created in subsequent years (until 2011). In doing so, I follow these steps. In some cases, multiple districts were created from a single parent district, and so I assign the treatment status of the parent district to all the new districts. In other cases, several districts were combined to form a new big district. Here, if all the parent districts had the same treatment status, I assign the same status to the new district. However, there are cases in which the DPEP treatment status of the parent districts differ. If so, if more than 50 percent of the population of the new district comes from parent districts of a certain treatment status, I assign this treatment status to the new district. I use analogous rules in assigning programme start years and district characteristics (such as the 1991 district female literacy rates) to the newly created districts.

[11] India is administratively split up first into states, which are then split up into districts. These are akin to counties in the US.

1.4.2 Empirical Methodology

The analysis here estimates the impact of DPEP on two different samples. The parent sample comprises of those who were of school going age (5–14 years) while DPEP was being implemented in their district; these are the direct beneficiaries of the program. The child sample consists of the children of the direct beneficiaries; these were indirect beneficiaries of DPEP, who started their schooling after DPEP had been phased out in 2005. I define the treatment group depending on when a certain district received the scheme. I use government archival records to infer the exact start year of the programme in each of the 271 treatment districts (discussed earlier).

My estimation strategy relies on two important sources of exogenous variation – 1. the district female literacy cut off of 39.2 percent, and 2. the spatial and temporal variation in the implementation of DPEP. With regards to the former, the programme was assigned on the rule that districts with female literacy below the national average rate (39.2 percent) were more likely to receive DPEP. As a result, when moving from the right (above) of the cutoff to the left (below), the probability of treatment receipt experiences a discontinuous increase. This is illustrated in figure 1.5, where I graph the probability of programme receipt against the female literacy rates of different districts. The figure illustrates that around the RD cutoff (39.2 percent) there is a large discontinuity in the probability of receiving the programme. This setup is a Fuzzy Regression Discontinuity (FRD) design, and allows me to estimate the impact of DPEP around the allocation cutoff.

The intuition here is to identify the effect of DPEP by comparing the outcomes of a subset of observations on either side of the RD cutoff ($\bar{x}$). This subset of observations lies within a neighborhood around the cutoff. Recent innovations in the field of RD estimation and inference make it possible to employ the underlying data to estimate the size of the neighborhood[12]. This is in contrast to the previously used methods, like 2SLS–IV, to compute the causal impacts (discussed later).

The neighborhood typically takes the following form: $[\bar{x} - h, \bar{x} + h]$, where h is the optimally determined bandwidth. There are two main data–driven approaches that can be used to calculate the optimal bandwidth – the Mean Squared Error (MSE) method and the Coverage Error Rate (CER). Although both approaches are semi–parametric in nature, and involve tradeoffs between efficiency and robustness, they differ in the optimality criterion used to calculate the bandwidth. In implementing these approaches, I specify several parameters to facilitate the bandwidth estimation. First, I select the kernel function that is to be used to determine the weight assigned to each observation. In my analysis I use a triangular kernel which puts higher weight on observations close to the RD cutoff and less weight on observations that are further away. I show that the results are robust to using an epanechnikov kernel. Second, I select the polynomial function form to used in the model estimation. To allow for more flexibility, I use a quadratic polynomial for the main results, but also use a linear function to show that the results are not sensitive to this change.

In estimating the impacts of DPEP on direct beneficiaries, I incorporate a series of control variables such as age of the individual and categorical variables for religion, caste, state and year of data collection. For the child level specifications, I control for child's age and gender, age of both parents, rainfall shocks

[12]These observations that are close to the cutoff on either side are similar on most characteristics, except their probability of receiving DPEP, something that I verify in the results section.

in–utero/birth year of the child (to proxy for environmental circumstances during this crucial period of growth)[13], and dummies for caste, religion, state and year of data collection. In all specifications, I cluster standard errors at the district level.

Imbens and Kalyanaraman, 2012 discuss the MSE approach and devise an asymptotically optimal procedure to estimate the bandwidth. Under this procedure, they assumed a squared error loss function and The formula used to determine the ideal/appropriate bandwidth under this procedure is:

$$h_{MSE} = C_{MSE}.n^{-1/(2p+3)} \tag{1.1}$$

where n is the sample size, p is the order of the polynomial chosen by the researcher. The constant C_{mse} depends on the kernel function, the polynomial form and the bias/variance of the estimator among other factors[14]. Calonico et al., 2017 discuss *robust–bias* corrections that make inference feasible with the MSE approach[15]. In my analysis, I report these robust–bias corrected standard errors along with the coefficient estimate.

[13] I use the same rainfall definition as used in the main analysis in Björkman-Nyqvist, 2013.

[14] This constant is unknown and needs to be estimated in order to ascertain the bandwidth (h_{mse}). Imbens and Kalyanaraman, 2012 propose a plug–in estimator that is based on a reference model to calculate an estimated value of the constant ($\hat{C}_{mse}$). This estimated value is then used to calculate the value of the bandwidth ($\hat{h}_{mse}$).

[15] Calonico et al., 2014 improved on the initial procedure suggested by Imbens and Kalyanaraman, 2012 by providing a bandwidth selector that has superior finite sample properties. In addition to being completely data driven and providing a mean squared error optimal bandwidth, this improved bandwidth selection procedure also has desirable small and large sample properties (Cattaneo and Vazquez-Bare, 2016). But, it has been shown that the standard errors of the RD estimate obtained from this procedure are not valid for inference. This is because the way the procedure balances between the bias and the variance makes inference logically inconsistent (for details refer to Calonico et al., 2014). This issue in these bandwidth selection procedures implies that the regular confidence interval that they produce cannot be used for inference. In the limiting case, where we assume a zero bias, the bandwidth size (h_{mse}) tends to infinity (since C_{mse} is inversely proportional to the bias).

However, Cattaneo and Vazquez-Bare, 2016 show that when inference is the goal of the estimation, then the MSE estimator and the associated robust–bias corrected confidence intervals may not be the preferred bandwidth selection approach. Their discussion demonstrates that the bandwidth value that reduces the Coverage Error (CE) of the confidence interval would be more appropriate. This is given by:

$$h_{CER} = C_{CER}.n^{-1/(p+3)} \tag{1.2}$$

where C_{CER} is a constant different from C_{MSE} and is estimated based on the underlying data. The confidence interval of the RD estimate based on this bandwidth (h_{CER}) has been shown to have demonstrably superior properties associated with inference[16][17].

These data–driven approaches are new to the literature and have not been used extensively in empirical applications.. RD analyses usually employ global polynomial approaches which tend to be subjective, not data driven and leads to larger bandwidths, While the global method works best when there is minimal misspecification bias (discussed in Gelman and Imbens, 2017), which is rare to achieve, the approach has appeal since it allow researchers to estimate causal impact with least squares estimation. I thus also estimate the impacts of DPEP with the 2SLS and present these results as a robustness check. I describe the global approach in Appendix A.

[16]In addition, the bandwidth which minimizes the Coverage Error (CE) is also always smaller than the bandwidth which minimizes the Mean Squared Error (MSE)That is, the number of observations used in the estimation using MSE is *larger* than (or equal to) the number of observations used in the estimation using the CE method.

[17] Cattaneo and Vazquez-Bare, 2016 note that owing to the large degree of variability in the point estimates, the RD coefficient from the CER procedure may not always be useful in empirical applications. Even so, I report the point estimates and the associated confidence intervals from these estimations. I primarily focus on the confidence intervals and discuss their relevance in assessing the statistical significance of the estimates.

1.4.3 RD Validity

For the RD design to be valid it is critical that individuals not be able to manipulate their treatment status by systematically positioning themselves on either side of the cutoff. If individuals can choose their own value of the running variable, then they can potentially decide whether or not to be a part of the treatment group. This would lead to non–random assignment to treatment, which would complicate the identification of the causal impact of the treatment. Such violations could occur in this case in two potential ways – if sub–national governments (at the state or district level) were able to choose their treatment status or if individuals were able to affect their treatment status through systematic migration. As a first step to establish this, I conduct the McCrary density test (as described in McCrary, 2008), the results of which are shown in figure 1.4. The graphs and the associated test statistic (p–value = 0.42) suggest that there is no discontinuity in the forcing variable (district female literacy rate) around the RD cutoff.

Additionally, it is unlikely that states/districts were able to manipulate their values of the running variable (district female literacy rate) since programme allocation was based on 1991 census data, which was collected at an earlier point of time by a central authority in India which is independent of state/district oversight. Additionally, the census data was collected in or before 1991, whereas the programme was announced in 1993. This meant that there was little chance that the states/districts knew about the programme when the census data was collected. Furthermore, it is highly likely that the states (or districts) had limited knowledge of the exact decision rule regarding the programme prior to DPEP implementation[18], more so because no other government programmes in the past (or since) appear to have been allocated based on the district female literacy rate.

While individuals could potentially have determined their treatment status through systematic migration across districts, I argue that this is unlikely in India and could not have been large enough to bias the estimates that I identify through my analysis. First, migration across districts in India in the 1990's was fairly low (Topalova, 2007). Second, the main reasons for migration in India are marriage and employment. Schooling choice (especially primary school) was not a major reason for migration in India, especially in rural India around the time DPEP was implemented (1993–2004). In terms of migration that is related to seeking enhanced education opportunities, most of it might be expected to be confined to the realms of higher education (high school and beyond). Since the DPEP programme mostly constructed primary, upper primary and secondary schools, the case for systematic migration affecting the composition of the treatment group seems weak.

1.5 Results

1.5.1 Discontinuity in Programme Receipt

As a first step, I show that there is a significant discontinuity in treatment assignment around the programme cutoff (the 1991 national average female literacy

[18] As decisions regarding programme placement were being made by the central government in conjunction with the World Bank and other donors

rate of 39.2 percent). Figure 1.5, which plots the probability of a district being part of the treatment group against the 1991 district female literacy rate, clearly illustrates that there is a significant difference in the probability of programme receipt around the RD cutoff. This implies that districts just below the cutoff were much more likely to be part of the DPEP treatment group as compared to districts that were just above the RD threshold. Despite their being a significant difference in probability of treatment assignment, it is possible that because of implementation issues this might not translate into differences in the actual number of schools constructed as a result of this programme. This is because it is possible for districts earmarked to receive the programme, due to a variety of reasons, to either not receive DPEP funding or be unable to use the funding properly. Therefore, in addition to showing discontinuity in programme assignment (as announced), it is also vital to establish that there is a significant break in the number of actual schools constructed during the DPEP implementation period in districts around the cutoff. I establish that in the next sub–section.

1.5.2 School Infrastructure (1993-2004)

I use district–level information to examine differences in school infrastructure between DPEP and non–DPEP regions at three time periods: in 1993 (Pre-DPEP), between 1993 & 2004 (DPEP years) and in 2005 (end of DPEP). To confirm whether the treatment–control differences observed above are indeed due to the DPEP policy and not due to pre–existing variations, I check whether any such discontinuities existed prior to program initiation in 1993. The results in panel A of Table 1.1 indicate that while treatment districts had marginally fewer number of schools in 1993 than control areas, the difference is statistically indistinguishable from zero. Panel B of Table 1.1 shows the impact of the DPEP programme on school construction during the DPEP years (1993–2004). The results indicate that an average DPEP district received almost 258 more government schools than a comparable non–DPEP district, a difference that is statistically significant. This difference persists when I examine total (public and private) schools and private schools separately, though the latter is not statistically significant. In panel B of Table 1.1, I also estimate the impact of DPEP on per capita schooling availability – the outcome I examine is the number of schools per 1000 individuals. The results indicate that there was a significant increase in the per capita availability of government schools (0.21 schools per 1000 population) and all schools (0.31 schools per 1000 population) in the treatment districts. I again fail to find any significant differences across treatment and control districts in the per capita availability of private schools.

To estimate the intergenerational impact of DPEP, we want the parent generation to benefit from enhanced school opportunities, but do not want their children to directly benefit from this policy. This would imply that there should not be significant differences in schooling access when the children start going to school – which is in the year 2005. I establish that this is the case in panel C of Table 1.1. The results indicate that there is no statistically significant difference in the total number and per capita (per 1000 population) government/private schools across the RD cutoff in the year 2005. This finding indicates that the results that I identify are likely to emerge solely due to the intergenerational effects of enhanced parental access to schooling.

1.5.3 School Quality (1993–2004)

Analogous to the analysis above, I test for differences in the quality of school infrastructure at three points in time: in 1993 (Pre–DPEP), between 1993 & 2004 (DPEP years) and in 2005 (end of DPEP). This would help understand if the programme had a significant impact on the underlying quality of schools in DPEP regions. I use DISE school–level census data from 2005 to examine several school quality measures – physical infrastructure (classrooms, toilets, electricity), teacher qualification, school oversight (inspection visits) and grants/incentives received by the school (funding received/spent). These arguably provide a comprehensive overview of the amenities/resources that a school possesses and is a wide enough array of indicators so as to encompass enough aspects of school quality. The results in Table 1.2 show that there were no differences in school quality at the start of the DPEP policy (in 1993)[19] and those constructed during the DPEP years (between 1993 & 2004)[20]. Additionally, I find that in the year 2005 there were no statistically significant differences on any of the quality indicators (Table 1.2). This alleviates concerns about the positive intergenerational learning effects being driven by superior school quality experienced by the children in the treatment group. It also further adds credence to the argument that any positive intergenerational effects of schooling observed in this context are due to enhanced schooling access of parents and

[19]Ideally, to do this one would have used data on these measures from the year 1993–94. Since such detailed data is unavailable for that time period, I use data from 2005 for schools that were constructed before 1993. This estimation would be valid if there were no systematic differences in upgrades/improvements in schools constructed before 1993 in districts around the programme cutoff. There is no reason to believe that this would be the case.

[20]Ideally, I would want to compare the schools built under the DPEP programme to other schools constructed in this time period to show that the DPEP schools were no different from the other schools. But the dataset does not identify the schools specifically built under this programme. So I compare all schools constructed in this period in districts around the cutoff. Given that DPEP was the flagship government programme for that period, it can be argued that any differences in school quality should be captured in this setup.

not due to improved school quality.

1.5.4 Effect on Direct Beneficiaries

I examine the impact that DPEP had on educational outcomes of the cohort of individuals who were of school–going age at the time of programme implementation. As discussed earlier, for the main estimation results I compare educational outcomes of DPEP direct beneficiaries, that is of people who were 14 years or below at the time of programme implementation across the RD cut-off. Table 1.3 presents estimates for the impact of DPEP on male and female direct beneficiaries separately. I find that the programme had a positive effect on enrolment in school, with both males and females experiencing an 8–10 percentage point increase. Male and female beneficiaries also had more years of education (0.75 – 0.9 years), were more likely to complete primary school (5 – 12 percentage points) and were nearly 9 percentage points more likely to be literate. These results indicate that those going to school in the immediate aftermath of DPEP initiation did attain higher schooling through the enhanced schooling access provided by the program. The effects are present for both genders.

1.5.5 Intergenerational Effects

I now present the main results of this analysis, the impact of the DPEP programme on the learning outcomes of the children of direct DPEP beneficiaries. As discussed in the empirical strategy section, I use estimators based on two different approaches – Mean Squared Error (MSE) and Coverage Error Rate

(CER). A child could have indirectly been exposed to the consequences of DPEP through either parent (mother or father) or both parents. In my analysis, I consider these cases separately.

Children – Mother was sole DPEP beneficiary

To compute treatment effects for this group, I compare test score outcomes for children whose mothers were part of the DPEP direct beneficiary group with those whose mothers were not, while in both cases the respective fathers were not impacted by DPEP due to their (they were too old to be of school–going age during DPEP years). I use the CER–optimal bandwidth estimator with a quadratic polynomial, and the test score outcomes are scored in a way such that zero means a failure to provide any correct answers and four indicates complete proficiency on the test. Column 1 of table 1.4 shows that when a woman benefitted from the DPEP programme (but her husband did not), her child's reading test score went up by 0.28 points, which is around 19 percent of the standard deviation (1.44) – which is also represented in figure 1.7. Since a one unit increase on the test implies an increase of one skill level, this can also be interpreted as an increase of a little more than one–fourth of a skill level. Given that an average child is close to being able to read a word (score = 2), the coefficient implies that DPEP was able to nudge the child of a typical female program beneficiary towards being able to read somewhere between a word and a sentence. The DPEP impact on math scores of 0.21 points (column 2 in table 1.4 and figure 1.6) would enable an average child to get closer to recognizing a two digit number (score = 2) rather than a one–digit number (score = 1). This is an improvement of almost 18 percent of the standard deviation (1.14) in math ability. I also find a positive effect of the programme on English reading ability by the children of treatment women – the RD coefficient is 0.10 points and it is significant at the five percent level (figure 1.8).

No studies have looked at the intergenerational learning impacts of a school construction programme, and hence there are no obvious studies to compare these results with. To provide some context, I look at other interventions that have aimed to improve learning outcomes in different countries. I find that the effect sizes I find are smaller than, but in line with, other studies from India (and other countries) that have looked at the impact of different school construction programmes (0.4 S.D.(σ) (Kazianga et al., 2013), 0.65 σ (Burde and Linden, 2013)) and other interventions on learning outcomes[21]: around 0.5 σ (Banerjee et al., 2007), 0.2 σ(Glewwe et al., 2009), around 0.2 σ (Duflo et al., 2012) and close to 0.3 σ (Muralidharan et al., 2016).

Children – Father was sole DPEP beneficiary

Akin to the analysis above, I examine outcomes for children whose fathers benefitted from the DPEP, but their mothers did not[22]. The results in Table 5 suggest that there was no statistically significant impact of the programme on this set of children. Although all the estimates (Table 1.5) are signed in the same way as the estimates for the children whose mothers were DPEP beneficiaries (Table 1.4), none of the impacts are statistically significant. This cannot be attributed to DPEP not benefiting male beneficiaries – recall, that male beneficiaries exposed to DPEP were found to have improved school attainment (Table 1.3). Neither

[21] σ here represents standard deviation

[22] Because the mothers were too young to be of school–going age during DPEP programme years

can this null result be due to small sample sizes within the bandwidth – the effective number of observations are more than 23,000 in each of the outcomes (except english score). Rather, it seems likely that mothers are able to use their enhanced schooling to improve the outcomes of their children, while fathers' ability to do the same seems restricted.

To further probe this result, I divide the children by gender and verify if it is the case that there are gender heterogeneities in the ability to transfer human capital benefits to children. The results in table 1.6 indicate two patterns – father beneficiaries are not able to benefit children of either gender. Second, daughters gain more from mother beneficiaries across all outcomes. These results are in line with other studies that find the role of mother to be vital in the human capital formation of children (Desai and Alva, 1998, Currie and Hyson, 1999, Currie and Madrian, 1999, Persico et al., 2004, Case et al., 2002, Behrman and Rosenzweig, 2004, Case et al., 2005, King et al., 2007, Güneş, 2015, Vollmer et al., 2016, Alderman and Headey, 2017). In addition, the higher impact on females (daughters) than on males (sons) is similar to the gender difference in the impact of school construction programme in Afghanistan, which was evaluated by Burde and Linden, 2013.

Both Parents Treatment

Table 1.7 looks at the sample of children with both parents benefitting from the DPEP program. The results are qualitatively similar to the results for the children who only had mothers exposed to DPEP school construction (Table 1.4). When these results are looked at in conjunction with those for the children with treated fathers, one may conclude that while enhanced maternal schooling through DPEP certainly mattered for child learning, paternal schooling might not have. As pointed out earlier, this is consistent with existing evidence in the development economics literature.

Putting RD–LATE in perspective

The RD estimation procedure leads to a Local Average Treatment Effect (LATE) of the causal impact of the programme, which is based on observations that are close to the RD cutoff. In this context, that would be people living in districts with female literacy rate (in 1991) close to 39.2 percent. Since this effect is a *local* estimate, this could mean that the coefficient might not be widely generalizable. Below, I present two different perspectives as why this criticism might not be as relevant in this case.

First, from a global perspective the results here would be informative about the impact of a similar school construction programme in other developing countries which are in the same stage of educational development as India was at the time DPEP was implemented. There are many countries with large populations, like Pakistan (200 million) and Ethiopia (100 million), that have a average female literacy rate around (or below) the RD cutoff (39.2 percent) of this study [23]. Most of these countries are concentrated in Central and West Africa and South Asia. Additionally, there are other countries in Africa and Asia that have higher rates of overall female literacy, but have large regions within them where educational indicators are similar to what they were in India at the start of the DPEP. The results here would also be potentially valuable for forming policies in these areas.

[23]Literacy data from UNICEF (2015) – link – accessed on 25 July 2018.

Second, I show that these results are valuable in the Indian context. If the districts around the cutoff were concentrated in one part of the country, then the results would not be valid for India as a whole. To verify whether this is the case, I plot on a graph the districts that are within a neighborhood of the allocation cutoff. I use the effects on intergenerational reading scores to illustrate my point. The impact of DPEP on the reading scores for the children of program beneficiaries is 0.28, which is based on more than 37,000 observations (out of a total of more than 480,000 observations) within a bandwidth of around 5 percent (running variable) on either side of the cutoff. While the bandwidth is fairly narrow, the number of observations is sizable, which is uncommon in RD applications of this nature. Additionally, these observations are not localized to a few districts around the cutoff – they come from 46 districts, spread across different parts of the country.

1.5.6 Falsification Checks

I conduct several checks to demonstrate that the results that I obtain are due to the DPEP school construction policy and are not due to any other factors. First, I verify whether the setup I use detects any impacts for groups that should have been un-affected by the DPEP programme – children who were 14 years or older at the time of programme implementation would have been too old to benefit from the school expansion under the DPEP, and should ideally show no programme impacts. The results in table 1.14 does show that this sub–sample of people show no effects of DPEP. Further, the children of these people should also not exhibit any effects of the programme. I verify this in Table 1.15 – I estimate the RD specification separately for children whose mothers just missed benefiting from the program (panels A and B) and for children whose fathers just missed being exposed (panels C and D). I find that in both cases, there are no DPEP impacts, indicating that the school construction programme had no statistically significant impact on the outcomes for the children of women (or men) who were likely to have left school or aged out of the school–going age range by the time DPEP was implemented in their districts. This further strengthens the main results.

I also use the same RD setup to estimate the impact on pre–determined or unrelated covariates, which have been determined independently of DPEP, and hence ideally should be unaffected by them. These include age of the mother, gender and age of the child and birth/current year rainfall shock[24]. The results from this exercise are shown in Figures 1.9 & 1.10. I plot the point estimates and their 90 percent confidence intervals that are estimated using different bandwidths and kernel functions (Triangle and Epanechnikov). The confidence interval of the point estimate of the impact of DPEP on these outcomes always consists of the zero value, showing that DPEP, did not shape outcomes unrelated to the programme.

1.5.7 Robustness Checks

Table 1.4 provides the main set of results for the case when the mother of the child is the sole DPEP beneficiary. In this section, I discuss the different robustness checks that I conduct. In each of them I alter different parts of the empiri-

[24]The data on rainfall comes from the University of Delaware dataset on precipitation and air temperature (Matsuura and Willmott, 2015). Any differences in current year rainfall around the RD cutoff can potentially be a confounder, but ideally this should not be the case. Therefore, I show this empirically to assuage any such concerns.

cal strategy used in table 1.4 – bandwidth estimation, the polynomial functional form and the kernel function used to assign weights to the observations around the cutoff [25].

First, I alter the approach to RD estimation – instead of using the CER approach (Table 1.4), I use an MSE based approach in table 1.8. In the MSE approach, the point estimates change – some increase while others decrease, but they always retain their statistical significance. Next, I check the sensitivity of the results to the type of kernel function chosen. While, I use a triangular kernel for the main results (Tables 1.4), I re–examine the results for the main outcomes with an epanechnikov kernel (in Table 1.9). The point estimates and the bandwidths do change marginally, but the point estimates mostly retain their sign and significance (except the enrolled outcome variable). I conduct another check on the same lines where I use a linear polynomial function (power = 1), instead of a quadratic polynomial that is used in the main results (Table 1.4). Like the change in kernel functional form, the inferences made from the main tables still mostly remain robust (Table 1.10).

In another check, I verify how the results are affected when I change the way in which the sample from the control districts is defined. While in the main analysis I use the statewise averages of DPEP districts to assign start dates to non–DPEP districts within the state, in Panel A of Table 1.11 I define the the control group sample using nationwide average start year of DPEP. Under this method I assign all control districts the nationwide average start date among DPEP districts. This method ignores the state (or regional) differences in DPEP implementation patterns across districts. The results (Panel A of Table 1.11)

[25]Tables for robustness check in the case of the father being the sole beneficiary of the programme show that the results are robust in that case. Tables are available from the author on request.

suggest that the main results are largely robust to these changes – the impact on all outcomes retain the right sign, while most of them remain statistically significant as well (except English score).

For the next test, I alter the way I define the cohort of individuals who would have directly benefitted from DPEP. In the main results, I define children who between 5 and 14 years during DPEP years to be the beneficiary group of DPEP school construction. It is plausible that since the majority of the schools constructed under this policy were primary (and upper–primary) schools, children younger than 14 years would have experienced most of the direct impact. Therefore, I re–estimate the results with a lower age cutoff of 12 years to define the cohort that might have been impacted by the programme[26]. This lower age cutoff is especially relevant for girls since they tend to drop out of schools at younger ages than boys due to a variety of reasons, chief among them being child marriage and onset of menarche. The effect of onset of menses on schooling attainment has been studied in developed (Burrows 1 and Johnson, 2005, Roberts et al., 2002, Joan and Zittel, 1998) and developing countries (Sommer, 2010). Other evidence finds that it may be the case that onset of menarche might lead to higher and earlier dropout from schools amongst girls (Adukia, 2017, Kirk and Sommer, 2006, Burgers, 2000, Fentiman et al., 1999). The results from this exercise are presented in panel B of Table 1.11 – the RD coefficients fall in magnitude, but retain their statistical significance for most outcomes. This implies that the results mostly remain stable when using this different (potentially stricter) definition of the treatment group.

Additionally, I check how the results change when I use the global polynomial approach to RD estimation, instead of a local polynomial approach. The

[26]This change would lead to the reduction in the size of the treatment group.

former uses the whole dataset to estimate the RD impact using a 2SLS–IV strategy. In Tables 1.12 and 1.13, I replicate the analysis from the main results using a global polynomial approach. I find that the overall pattern of results does not change in cases when the mother was exposed to the DPEP programme, it had a positive impact on children's reading, math and English test scores while reducing the chances of not being able to answer any questions on these tests.

1.6 Mechanisms

There are potentially multiple pathways through which a school construction programme (like DPEP) could shape intergenerational learning outcomes in a developing country like India. While results that I present earlier (Table 1.3) show that individuals who directly benefitted from DPEP were able to increase their school attainment, we don't know what it was about this schooling that enabled them to positively impact their children's learning outcomes. I now examine several potential pathways that could have been responsible for the observed intergenerational effects. Given that female DPEP beneficiaries appear to be most able to use their education to shape their children's lives, here I focus on the women beneficiaries and the sample of children who had mothers, but not fathers, who benefitted from the program.

1.6.1 Educational Investments

It is possible that highly educated parents might invest more in their children's education than less educated parents, and I use IHDS data to examine whether DPEP's intergenerational education effects could have been transmitted through such a channel. One way that parents can do this is to enroll their children in potentially higher quality schools, which in the Indian context could mean private schools[27]. I estimate whether DPEP programme exposure had an impact on parents' choice between private and government schools and find that there was no statistically significant impact on private school enrolment (Table 1.17). Higher investment in children could also take the form of greater schooling related expenditures (example – on books, tuition etc.). Results in 1.17 suggest that DPEP beneficiaries allocate more resources towards the payment of school fees, and the purchase of books and uniforms for children, while children of programme beneficiaries spent around two more hours doing homework than comparable children of non-beneficiaries, which might be due to the increased supervision by their mothers. This is similar to the results found by Andrabi et al., 2012 in a similar context (Pakistan). Therefore, there is some evidence that DPEP's intergenerational impacts might have been mediated through higher parental investments in children's education.

1.6.2 Health of Direct Beneficiaries

Extensive research shows that health in infancy (especially birthweight) has a significant impact on later life outcomes for children (Black et al., 2007, Oreopoulos et al., 2008, Royer, 2009, Bharadwaj et al., 2010). Additionally, it is well established that mother's health (and health behaviors) are key determinants of the health and well–being of her children (Ahlburg, 1998, Coneus and

[27]Evidence from India shows that private school attendance in India leads to large improvements on English test scores and a moderate impact on mathematics and vernacular test scores (**?**). This is despite the fact that private schools pay teachers lesser and spend less per pupil than government schools (Desai et al., 2009, Kingdon, 2007, Muralidharan and Kremer, 2006).

Spiess, 2012, Bhalotra and Rawlings, 2013, Yan, 2015). Therefore, it is plausible that DPEP's positive effect on female education had a knock–on effect on their health, and the latter led to higher well–being of children. Using data from the Annual Health Survey (2012–13) I test this hypothesis. I find that DPEP female beneficiaries indeed had better health as adults (as measured by BMI and hemoglobin – more details in Sunder, 2018b). This in line with findings from other studies that provide evidence on the positive impact of women's education on their own health (Grossman, 2015, Grépin and Bharadwaj, 2015, Agüero and Bharadwaj, 2014, Lundborg, 2013, Amin et al., 2013, Silles, 2009, Currie and Moretti, 2003). In addition, I find that DPEP had a beneficial impact on female contraceptive usage – which might reduce unwanted fertility in the high fertility context of rural India (table 1.16), which in turn might foster higher human capital of children (Kugler and Kumar, 2017), which resonates with findings from other studies (Johnston et al., 2015, Andalón et al., 2014).

1.6.3 Child Care Investments

A child's human capital is significantly impacted by in–utero and early life conditions, or what is known as the first 1000 days of life (Almond and Currie, 2011a and Currie and Vogl, 2013 provide good reviews of this literature). Did DPEP exposure lead to higher usage of Ante Natal Care (ANC) and Post Natal Care (PNC) by beneficiaries when pregnant? Such services would have lead to better outcomes for children as well as for the women themselves (Paudel et al., 2014, Onasoga et al., 2012, Simkhada et al., 2008, Kerber et al., 2007). In Table 1.16, I find that the women DPEP beneficiaries were more likely to make at least one ANC visit (7–11 percent), make more ANC visits in total (0.2–0.3 visits), obtain Iron and Folic Acid (IFA)[28] (3–10 percent), deliver in a facility (6–8 percent) and make at least one PNC visit (6 percent). Based on these findings, it is clear that women who were impacted by DPEP are more likely to receive care during and after their pregnancy, which arguably could have led to the enhanced child level human capital effects later in their lives.

1.6.4 Marriage Outcomes & Bargaining Power

Previous studies have shown that women who stayed enrolled longer in schools, tended to marry at a higher age, and consequently experienced improved outcomes in adulthood such as enhanced bargaining power (Lundberg and Pollak, 1993, Field and Ambrus, 2008b, Duflo, 2012, Samarakoon and Parinduri, 2015, Crandall et al., 2016, Sunder, 2018a, Yount et al., 2018). As the next set of potential mechanisms, I test whether DPEP impacted such outcomes (using IHDS data). The results in Table 1.16 show that women beneficiaries married about half a year later and had their first birth 0.25 years later than women in the control group. I also examine the social status of female beneficiaries in the households they married into, where I find that the DPEP women report a higher likelihood of participating in decisions related to their children and household meals (Table 1.17). These women are also less likely to say that physical violence (by husbands) against wives is justified. It thus seems like the women who benefitted from DPEP have higher bargaining power within the households and might be able to shape their children's outcomes more effectively than non–beneficiaries (Yoong, 2012, Bono et al., 2016).

[28]This is an especially important outcome which addresses Iron Deficiency Anemia among pregnant women – a major health concern in the context of India (see Rai et al., 2018 for a recent discussion on this)

1.7 Conclusion

In this paper, I use the geographic and temporal variation in the implementation of a national–level school construction programme to conduct a Regression Discontinuity analysis to estimate its impact on intergenerational learning outcomes. The timing of implementation varied across the 271 treatment districts, which I account for using detailed government archival data. I first demonstrate that the programme engendered increased access to schooling in treatment districts during the period that DPEP was in operation (1993 to 2004). I find that individuals exposed to the program (of both genders) were more likely to be literate and complete more years of education than comparable individuals in districts that did not receive the program. Further, I find that children of female DPEP beneficiaries experienced positive effects on vernacular reading, math and English test scores. In contrast, male beneficiaries were unable to transfer any benefits to their children.

I conduct multiple robustness checks to establish the internal validity of the results of my analysis. I also validate the results through a placebo test – I show that individuals too old to benefit from DPEP (and their children) show no effects of the programme. In demonstrating the potential generalizability of these findings, I note two key points – first, that although the estimates are based on comparing individuals in districts close to the program cutoff, the sample consists of individuals from different parts of the country. This makes the results nationally relevant. Second, there are many countries in Africa (like Ethiopia and Ivory Coast) and South Asia (Pakistan and Afghanistan) that have female literacy levels close to or lower than the RD cutoff (39.2 percent) in this study. Therefore, I argue that even though I am able to identify the Local Average

Treatment Effects of DPEP, I am able to do so at a point in the female literacy distribution which approximates those prevailing in many developing countries. Therefore, the results from this analysis can possibly inform the education policies in these parts of the world.

In this analysis, I also explore the potential mechanisms through which the intergenerational impacts of the school construction could have been mediated. I find that women who were able to enhance their educational attainment through DPEP had better health (BMI) and superior health behavior in terms of contraceptive usage, pre–natal care and post–natal care as compared to non–beneficiary women. I also find DPEP's female beneficiaries married later, and had higher bargaining power in their marital households. All these factors could have enabled women to allocate greater resources towards their children's welfare. In fact, I do find that the children of these women benefitted from higher spending on school fees and uniforms/books.

Cognitive development and learning in childhood has an important bearing on later life outcomes and policy needs to focus on ways to enhance these outcomes. The bulk of the literature has focused on school based reforms to improve learning outcomes (Kremer et al., 2013, Muralidharan, 2013). Through this analysis I demonstrate that parents, particularly mothers, play an important role in shaping their children's ability to learn. Some interventions have been found to increase parental investment in children include providing parents with accurate info on returns to schooling (Bettinger and Slonim, 2007, Jensen, 2010, Levitt et al., 2011). Additionally, as the results of this analysis (and Andrabi et al., 2012, Banerji et al., 2017) show, improving the skill set of mothers could go a long way in boosting child performance on cognitive tests. Therefore, there is a need for policy action that targets parents to potentially increase educational investment in their children (Houtenville and Conway, 2008, Andrabi et al., 2015, Bergman, 2015). These reforms should complement, and not substitute, the school–based reforms aimed at improving child learning.

1.8 Tables

Table 1.1: Impact of DPEP on School Construction

Panel A: All Schools in 1993 (DPEP Start Year)

	Total Number of Schools			Schools per 1000 Population		
	All Schools	Government	Private	All Schools	Government	Private
RD Estimate	-143.7	-104.2	-50.77	-0.05	-0.04	-0.01
S.E. (coef)	343.9	301.6	83.2	0.28	0.49	0.03
Total Obs.	495	495	495	488	488	488

Panel B: Schools Constructed Between 1993 & 2005 (During DPEP Years)

	Total Number of Schools			Schools per 1000 Population		
	All Schools	Government	Private	All Schools	Government	Private
RD Estimate	413.8**	258.1**	148.5	0.31**	0.21**	0.08
S.E. (coef)	173.3	125.6	103	0.14	0.11	0.07
Total Obs.	495	495	495	488	488	488

Panel C: All Schools in 2005 (DPEP End Year)

	Total Number of Schools			Schools per 1000 Population		
	All Schools	Government	Private	All Schools	Government	Private
RD Estimate	270.1	153.9	107.7	0.28	0.19	0.09
S.E. (coef)	335.9	359.4	128.4	0.40	0.37	0.06
Total Obs.	495	495	495	488	488	488

Based on author's calculations using the District Information on System of Education (DISE) district level data for the year 2005. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. The standard errors are robust–bias corrected and are clustered at the district level.

LINK TO RESULTS SECTION

Table 1.2: Impact of DPEP on School Quality

School Infrastructure

	# Classrooms			Any Common Toilet			Any Girls Toilet			Any Electricity		
	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005
RD Estimate	5.06	3.6	6.5	0.68	-0.35	0.46	-0.90	-0.6	0.03	0.78	-0.08	0.58
S.E. (coef)	62.9	3.1	16.1	3.9	0.4	2.26	4.3	0.6	2.5	6.6	0.4	2.3
Total Obs.	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894

Teacher Characteristics

	# Male Teachers			# Female Teachers			# Graduate Teachers			Professional Qual.		
	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005
RD Estimate	-2.2	-0.79	-3.22	6.12	3.05	5.33	4.79	1.24	3.44	4.99	2.06	6.74
S.E. (coef)	15.1	1.22	9.2	18.8	3.1	10.1	19.2	1.9	7.5	21.1	1.84	11.2
Total Obs.	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894

School Oversight

	Distance – Block (kms)			Distance – Cluster (kms)			# Visits – Block			# Visits – Cluster		
	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005
RD Estimate	-20.8	-7.54	-9.22	7.62	0.15	8.04	9.09	2.62	6.15	-10.76	-1.41	-8.49
S.E. (coef)	49.5	9.4	7.9	12.2	3.08	36.6	58.6	2.37	12	22.4	3.1	27.6
Total Obs.	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894

Grants & Incentives

	Devt. Grant – Received ('000 Rs.)			Devt. Grant – Spent ('000 Rs.)			TLM Grant – Received ('000 Rs.)			TLM Grant – Spent ('000 Rs.)		
	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005	In 1993	1993-2004	In 2005
RD Estimate	-0.68	-0.11	0.34	-0.96	-0.11	0.23	-0.05	0.16	0.34	0.11	0.08	0.26
S.E. (coef)	8.9	0.2	0.9	8.6	0.16	0.8	1.68	0.12	0.5	1.46	0.05	0.3
Total Obs.	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894	726,494	291,280	1,017,894

Based on author's calculations using the District Information on System of Education (DISE) district level data for the year 2005. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. The standard errors are robust–bias corrected and are clustered at the district level. *TLM Grant* refers to grants received under the Total Literacy Mission. The **triangular kernel** with **local polynomial of order 2** is used to construct the point estimates. Estimates are based on author's calculations using the individual school level data from DISE (2005).

LINK TO RESULTS SECTION

Table 1.3: Direct Beneficiary Impacts

District Level Household & Facility Survey (DLHS) Round 3 (2007-08)								
	Panel A: Female Sample				Panel B: Male Sample			
	Ever School	Highest Grade	Completed Primary	Literate	Ever School	Highest Grade	Completed Primary	Literate
RD Estimate	0.10***	0.84***	0.12***	0.09***	0.09***	0.90***	0.05***	0.08***
S.E. (coef)	0.02	0.15	0.02	0.02	0.02	0.24	0.02	0.02
Total Obs.	110,543	110,517	110,517	110,212	92,098	90,759	90,759	90,756

District Level Household & Facility Survey (DLHS) Round 4 (2011-12)								
	Panel A: Female Sample				**Panel B: Male Sample**			
	Ever School	Highest Grade	Completed Primary	Literate	Ever School	Highest Grade	Completed Primary	Literate
RD Estimate	0.08***	0.75***	0.11***	-	0.08***	0.78***	0.05***	-
S.E. (coef)	0.01	0.17	0.03	-	0.02	0.21	0.01	-
Total Obs.	101,513	101,233	101,233	-	90,976	90,116	90,116	-

Note: The sample for this table consists of people who were below the age of 14 years at the time of programme implementation (from government archives data) in the treatment districts (DLHS data Rounds 3 & 4). The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD impact point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector bandwidth selector. All the specifications control for the age of the individual and categorical variables for caste, religion, state and year of data collection. Standard errors are robust and clustered at the district level.

LINK TO RESULTS SECTION

Table 1.4: Impact on Children when mother is sole DPEP beneficiary – CER Optimal

	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.28***	0.21**	0.10**	0.04*	0.06***
S.E. (coef)	0.11	0.10	0.05	0.024	0.018
Total Obs.	488,862	487,037	253,172	472,338	526,087
Bandwidth	2.4	1.5	1.5	0.9	3.9
Effective Obs.	37,203	28,240	13,159	25,604	51,154
Mean (Y)	1.93	1.70	1.5	0.85	0.90
S.E. (Y)	1.44	1.14	1.1	0.20	0.15

Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. The sample consists of children satisfying two criterion – likely started school after the year 2005 (DPEP end year) and that their mother was below the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth for each outcome is indicated.

LINK TO RESULTS SECTION

Table 1.5: Impact on Children when father is sole DPEP beneficiary – CER Optimal

	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.28	0.28	0.22	0.07	0.002
S.E. (coef)	0.22	0.27	0.16	0.07	0.04
Total Obs.	142,217	141,821	73,713	129,773	131,605
Bandwidth	6.1	4.7	6.1	4.9	5
Effective Obs.	32,672	26,172	16,680	23,662	23,936
Mean (Y)	2.39	2.16	2.08	0.84	0.92
S.E. (Y)	1.44	1.27	1.45	0.29	0.09

Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. The sample consists of children satisfying two criterion – likely started school after the year 2005 (DPEP end year) and that their father was below the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The standard errors are robust and are clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth for each outcome is indicated.

LINK TO RESULTS SECTION

Table 1.6: Impact of DPEP on child outcomes – Gender Heterogeneity

	Mother to Daughter			**Mother to Son**		
	Read Score	Math Score	English Score	Read Score	Math Score	English Score
RD Estimate	0.29***	0.27**	0.13***	0.26***	0.23*	0.11**
S.E. (coef)	0.10	0.13	0.05	0.10	0.12	0.04
Total Obs.	230,101	229,266	119,616	253,913	252,977	133,556
	Father to Daughter			**Father to Son**		
	Read Score	Math Score	English Score	Read Score	Math Score	English Score
RD Estimate	0.29	0.39	0.30	0.26	0.37	0.24
S.E. (coef)	0.23	0.32	0.24	0.23	0.34	0.16
Total Obs.	68,264	68,074	34,645	73,953	73,747	39,068

Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. The sample consists of children satisfying two criterion – likely started school after the year 2005 (DPEP end year) and that their father was below the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth for each outcome is indicated.

LINK TO RESULTS SECTION

Table 1.7: Impact on Children when both parents are DPEP beneficiaries – CER Optimal

	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.22***	0.22**	0.12**	0.03**	0.05*
S.E. (coef)	0.08	0.11	0.06	0.015	0.027
Total Obs.	137,980	137,611	71,325	125,660	127,408
Bandwidth	3.6	3.6	3.9	2.1	3.7
Effective Obs.	11,959	11,742	6,575	7,349	12,021
Mean (Y)	1.98	1.78	1.55	0.88	0.91
S.E. (Y)	1.46	1.16	1.12	0.21	0.14

Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. The sample consists of children satisfying two criterion – likely started school after the year 2005 (DPEP end year) and that both their mother and father were below the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The standard errors are robust and clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth for each outcome is indicated.

LINK TO RESULTS SECTION

Table 1.8: Impact on Children when mother is sole DPEP Beneficiary – MSE Optimal

	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.22***	0.19**	0.11**	0.05**	0.08***
S.E. (coef)	0.08	0.10	0.05	0.025	0.022
Total Obs.	488,862	487,037	253,172	472,338	526,087
Bandwidth	5.4	3.3	2.8	8.9	7.5
Effective Obs.	73,180	46,664	21,406	102,179	80,078
Mean (Y)	1.97	1.70	1.53	0.84	0.9
S.E. (Y)	1.44	1.16	1.12	0.22	0.15

Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. The sample consists of children satisfying two criterion – likely started school after the year 2005 (DPEP end year) and that their mother was below the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common MSE–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The standard errors are robust–bias corrected and are clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth for each outcome is indicated.

LINK TO RESULTS SECTION

Table 1.9: Robustness Check – Epanechnikov Kernel

CER Optimal	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.29***	0.22**	0.10	0.04*	0.04
S.E. (coef)	0.11	0.10	0.08	0.023	0.04
Total Obs.	488,862	487,037	253,172	472,338	526,087
Bandwidth	1.5	2.5	2.6	1.1	4.6
Effective Obs.	13,791	25,660	32,370	28,203	26,469
MSE Optimal	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.23***	0.24*	0.10**	0.05*	0.07
S.E. (coef)	0.09	0.14	0.05	0.028	0.08
Total Obs.	488,862	487,037	253,172	472,338	526,087
Bandwidth	1.6	3.6	3.6	8.4	6.1
Effective Obs.	26,981	56,560	56,381	94,230	139,999

Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. The sample consists of children satisfying two criterion – likely started school after the year 2005 (DPEP end year) and that their mother was below the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD point estimates are constructed using the epanechnikov kernel, local polynomial of order 2 and with one common CER–optimal & MSE–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The standard errors are robust–bias corrected and are clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth for each outcome is indicated.

LINK TO RESULTS SECTION

Table 1.10: Robustness Check – Linear Polynomial

CER Optimal	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.28***	0.27**	0.08*	0.06**	0.09
S.E. (coef)	0.12	0.13	0.102	0.04	0.06
Total Obs.	488,862	487,037	253,172	472,338	526,087
Bandwidth	0.92	0.96	3.1	0.92	0.98
Effective Obs.	7,140	19,780	25,603	27,203	14,756
MSE Optimal	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.33***	0.29**	0.09*	0.04***	0.06
S.E. (coef)	0.15	0.14	0.05	0.01	0.05
Total Obs.	488,862	487,037	253,172	472,338	526,087
Bandwidth	1.5	1.8	2.7	8.2	1.84
Effective Obs.	14,642	32,451	37,669	91,661	30,274

Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. The sample consists of children satisfying two criterion – likely started school after the year 2005 (DPEP end year) and that their mother was below the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 1 and with one common CER–optimal & MSE–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The standard errors are robust–bias corrected and are clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth for each outcome is indicated.

LINK TO RESULTS SECTION

Table 1.11: Robustness Checks – Different Start Years – CER RD

	Panel A: National Avg. Start			Panel B: Control = Age Cutoff = 12 yrs		
	Read Score	Math Score	English Score	Read Score	Math Score	English Score
RD Estimate	0.26***	0.20***	0.12	0.25***	0.26***	0.11***
S.E. (coef)	0.09	0.07	0.23	0.09	0.09	0.04
Total Obs.	486,264	484,428	247,489	309,775	310,074	164,351

Note: The score variables run from 0–4, whereas the other outcomes are categorical variables. The sample for this table consists of children born in or after the year 2000 to parents who were both below the age of 14 years at the time of programme implementation (from government archives data) in the treatment districts (ASER data, 2007–2014). The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD impact point estimates are constructed with **triangular kernel**, **local polynomial of order 2** and with **one common CER–optimal bandwidth selector** bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. The bandwidth is expressed in terms of percentage of the running variable (district female literacy rate in 1991). The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The mean and standard deviation of the dependent variable within the bandwidth of that particular outcome are included in the table.

LINK TO RESULTS SECTION

Table 1.12: Robustness Check – Full Sample Regression – Exact Timing

	Read Score	Math Score	English Score	GFA	Enrolled
RD Estimate	0.19***	0.10***	0.09***	0.024***	0.002
S.E. (coef)	0.06	0.035	0.03	0.004	0.002
Total Obs.	488,862	487,037	253,172	472,338	526,087

Note: The sample consists of children born in or after 2000 to mothers who were below 14 years of age at the time of implementation of the DPEP programme in their district. The score variables run from 0–4, whereas the other outcomes are categorical variables. GFA refers to a Grade–for–Age measure. This programme implementation timing is derived from detailed government archives that describe the exact process of programme implementation. The corresponding population in the control districts is identified on the basis of the average start date in treatment districts within the same state. All specifications control for a quadratic polynomial of the running variable, the child's age, ages of both parent, rainfall shocks in–utero/birth year of the child, and dummies for caste, religion, state and year of data collection. The standard errors are robust and are clustered at the district level.

LINK TO RESULTS SECTION

Table 1.13: Robustness Check – Full Sample Regression – Different Start Years

	Panel A: Control = National Avg. Start			Panel B: Age Cutoff = 12 yrs		
	Read Score	Math Score	English Score	Read Score	Math Score	English Score
RD Estimate	0.26***	0.20***	0.14**	0.23***	0.16***	0.14***
S.E. (coef)	0.07	0.06	0.06	0.09	0.06	0.04
Total Obs.	486,264	484,428	247,489	309,775	310,074	164,351

Note: The score variables run from 0–4. The sample consists of children born in or after 2000 to mothers who were below 14 years of age in a particular year. This year is the estimated start date of the programme using the DISE dataset – the year with the maximum year on year rate of growth of schools in a particular district after the implementation of the DPEP programme. All specifications control for a quadratic polynomial of the running variable, the child's age, ages of both parent, rainfall shocks in–utero/birth year of the child, and dummies for caste, religion, state and year of data collection. Standard errors are robust and clustered at district level.

LINK TO RESULTS SECTION

Table 1.14: Falsification – Direct Beneficiary Impacts

District Level Household & Facility Survey (DLHS) Round 3 (2007-08)								
	Panel A: Female Sample				Panel B: Male Sample			
	Ever School	Highest Grade	Completed Primary	Literate	Ever School	Highest Grade	Completed Primary	Literate
RD Estimate	-0.08	-0.54	-0.08	-0.02	-0.07	-0.34	-0.06	-0.01
S.E. (coef)	0.06	0.48	0.05	0.02	0.06	0.26	0.04	0.02
Total Obs.	271,978	271,940	271,940	269,320	189,231	188,764	188,764	188,223

District Level Household & Facility Survey (DLHS) Round 4 (2011-12)								
	Panel A: Female Sample				**Panel B: Male Sample**			
	Ever School	Highest Grade	Completed Primary	Literate	Ever School	Highest Grade	Completed Primary	Literate
RD Estimate	0.15	0.35	0.11	-	0.11	0.38	0.09	-
S.E. (coef)	0.12	0.22	0.08	-	0.09	0.27	0.06	-
Total Obs.	116,593	113,760	113,760	-	101,982	101,124	101,124	-

Note: The score variables run from 0–4. The sample for this table consists of people who were below the age of 14 years at the time of programme implementation (from government archives data) in the treatment districts (DLHS data Rounds 3 & 4). The starting year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD impact point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector bandwidth selector. All the specifications control for the age of the individual and categorical variables for caste, religion, state and year of data collection. Standard errors are robust and clustered at the district level.

LINK TO RESULTS SECTION

Table 1.15: Falsification Check – DPEP impact on children of non–beneficiaries

	Panel A: Non-Beneficiary Mother – CER			Panel B: Non-Beneficiary Mother – MSE		
	Read Score	Math Score	English Score	Read Score	Math Score	English Score
RD Estimate	-0.14	-0.07	-0.13	-0.09	-0.04	-0.24
S.E. (coef)	0.47	0.46	0.95	0.37	0.49	0.88
Total Obs.	737,551	733,786	382,316	737,551	733,786	382,316
Bandwidth	7.5	6.1	7.6	10.2	8.9	10.3
Effective Obs.	161,474	131,288	81,412	266,237	183,369	131,179

	Panel C: Non-Beneficiary Father – CER			Panel D: Non-Beneficiary Father – MSE		
	Read Score	Math Score	English Score	Read Score	Math Score	English Score
RD Estimate	-0.11	-0.16	-0.29	-0.13	-0.29	-0.22
S.E. (coef)	0.47	1.75	0.32	0.12	1.17	2.2
Total Obs.	584,862	582,528	266,994	584,862	582,528	266,994
Bandwidth	7.8	5.3	5.1	7.7	7.5	7.2
Effective Obs.	154,024	99.949	45,407	148,117	147,485	65,504

Author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The sample consists of children satisfying two criteria – likely started school after the year 2005 (DPEP end year) and that their mother (or father) was above the age of 14 years at the time of DPEP implementation (start year from government archives data). The start year of the programme for control districts is the state average starting year of treatment districts within the same state. The RD estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal & MSE–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in-utero/birth year of the child and dummy variables for state and year of data collection, and standard errors are robust and clustered at the district level. The bandwidth is expressed in terms of the running variable – district female literacy rate in 1991. The effective number of observations indicates the number of observations that lie within the bandwidths indicated in the table – these are different from the the full sample sizes which are also indicated in the table. The score variables run from 0–4, whereas the other outcomes are categorical variables. The mean and standard deviation of the dependent variable within the bandwidth is indicated.

LINK TO RESULTS SECTION

Table 1.16: Potential Mechanisms – Woman (Parent) Level

	Marriage Age		Age at First Birth		Contraceptive Use		Any ANC	
	DLHS-3	DLHS-4	DLHS-3	DLHS-4	DLHS-3	DLHS-4	DLHS-3	DLHS-4
RD Estimate	7.53***	6.62***	3.72***	3.01***	0.06***	0.07***	0.07***	0.11***
S.E. (coef)	0.52	0.68	0.42	0.6	0.02	0.03	0.02	0.03
Total Obs.	110,564	89,773	73,628	72,712	72,775	55,033	64,276	47,993

	# ANC Visits		IFA Taken		Delivery – Formal		Any PNC	
	DLHS-3	DLHS-4	DLHS-3	DLHS-4	DLHS-3	DLHS-4	DLHS-3	DLHS-4
RD Estimate	0.22**	0.31	0.032**	0.098**	0.063***	0.083***	0.058**	-
S.E. (coef)	0.12	0.03	0.014	0.038	0.021	0.027	0.024	-
Total Obs.	48,482	39,497	41,228	43,041	41,212	47,560	64,274	-

Source: Based on authors calculations using the District Level Household and Facility Survey (DLHS) data from Rounds 3 (2007–08) and Round 4 (2011–12). The sample consists of women who were below 14 years of age at the time of implementation of the DPEP programme in their district. This programme implementation timing is derived from detailed government archives that describe the exact process of programme implementation. The corresponding population in the control districts is identified on the basis of the average start date in treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the same set of variables as the main specifications. The standard errors are robust–bias corrected and are clustered at the district level. Marriage age refers to the age at marriage (in months), Age at first birth refers to age when the woman had her first child (In months), Contraceptive use is a dummy that takes a value of one if the women reported using contraceptives, *Any ANC* is a categorical variable that takes a value of one if the woman accessed any ANC facilities during the last pregnancy, # ANC visits refers to the number of ANC visits made during the last pregnancy, *IFA Taken* is a categorical variable that takes a value of one if the woman reported taking IFA tablets during the last pregnancy, *Delivery–Formal* is a categorical that takes a value of one if the woman reported giving birth in a formal health facility and *Any PNC* refers to a dummy that takes a value of one if the woman reported using any Post Natal Care (PNC) facilities.

LINK TO MECHANISMS SECTION

Table 1.17: Potential Mechanisms – Woman (Parent) & Child level

	Decision–Child	Decision–Cook	Decision–Purchases	Beat–Bad Cook	Beat–Neglect House
RD Estimate	0.08***	0.14***	0.08***	-0.09***	-0.05***
S.E. (coef)	0.02	0.053	0.03	0.02	0.01
Total Obs.	13,159	13,159	13,159	13,114	13,114

	School Fees	Uniform/Books	Tuition Fees	Private School	Homework Hours
RD Estimate	781.8*	706.8**	-15.6	0.05	2.06*
S.E. (coef)	426.3	328.6	312.8	0.05	1.16
Total Obs.	4,545	4,545	4,545	4,545	4,545

Source: Based on authors calculations using the Indian Human Development Survey (IHDS) 2005 round. The sample consists of women who were below 14 years of age at the time of implementation of the DPEP programme in their district. This programme implementation timing is derived from detailed government archives that describe the exact process of programme implementation. The corresponding population in the control districts is identified on the basis of the average start date in treatment districts within the same state. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the same set of variables as the main specifications. The standard errors are robust–bias corrected and are clustered at the district level. Decision–Cook and Decision–Purchases are dummy variables that take a value of one if the woman reported being involved in decisions related to cooking and purchases made in the household. *Beat–Bad Cook* and *Beat–Neglect House* are categorical variables that take a value of one if woman reported that it was common for women in their community to be beaten up in cases when she was a bad cook or neglected household work respectively. School Fees, Uniform/Books and Tuition fees refer to variables that measure the expenditure on these categories made on the woman's children. Private School is a dummy that takes a value of one if the child goes to private school, and homework hours refers to the amount of time the child spent doing homework.

LINK TO MECHANISMS SECTION

1.9 Figures

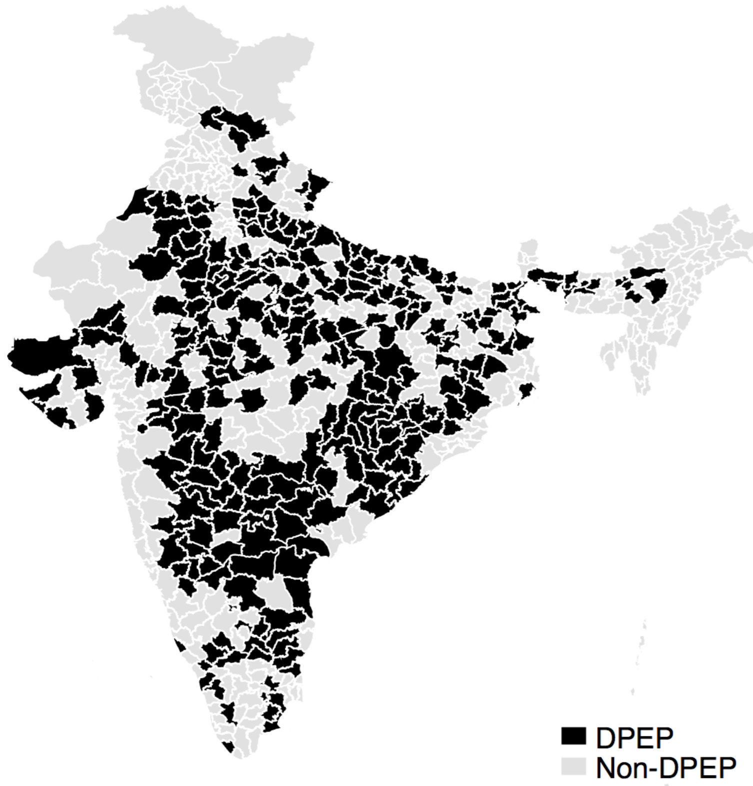

Figure 1.1: Sample of districts that received the DPEP programme
LINK TO BACKGROUND SECTION

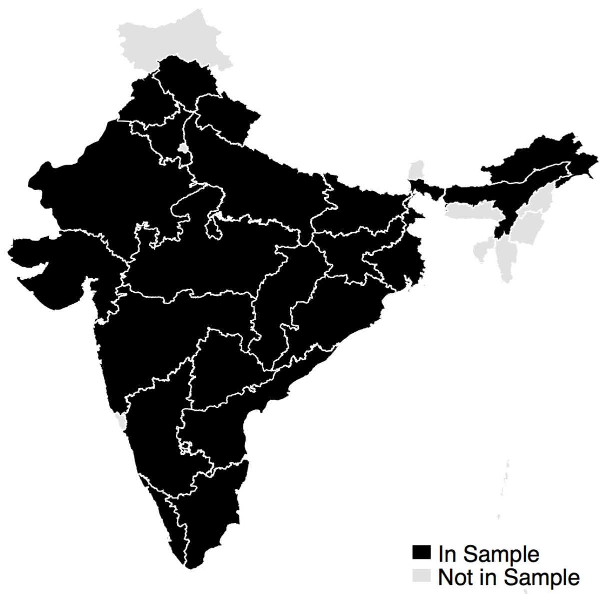

Figure 1.2: States included in my sample (ASER DATA).
LINK TO DATA SECTION

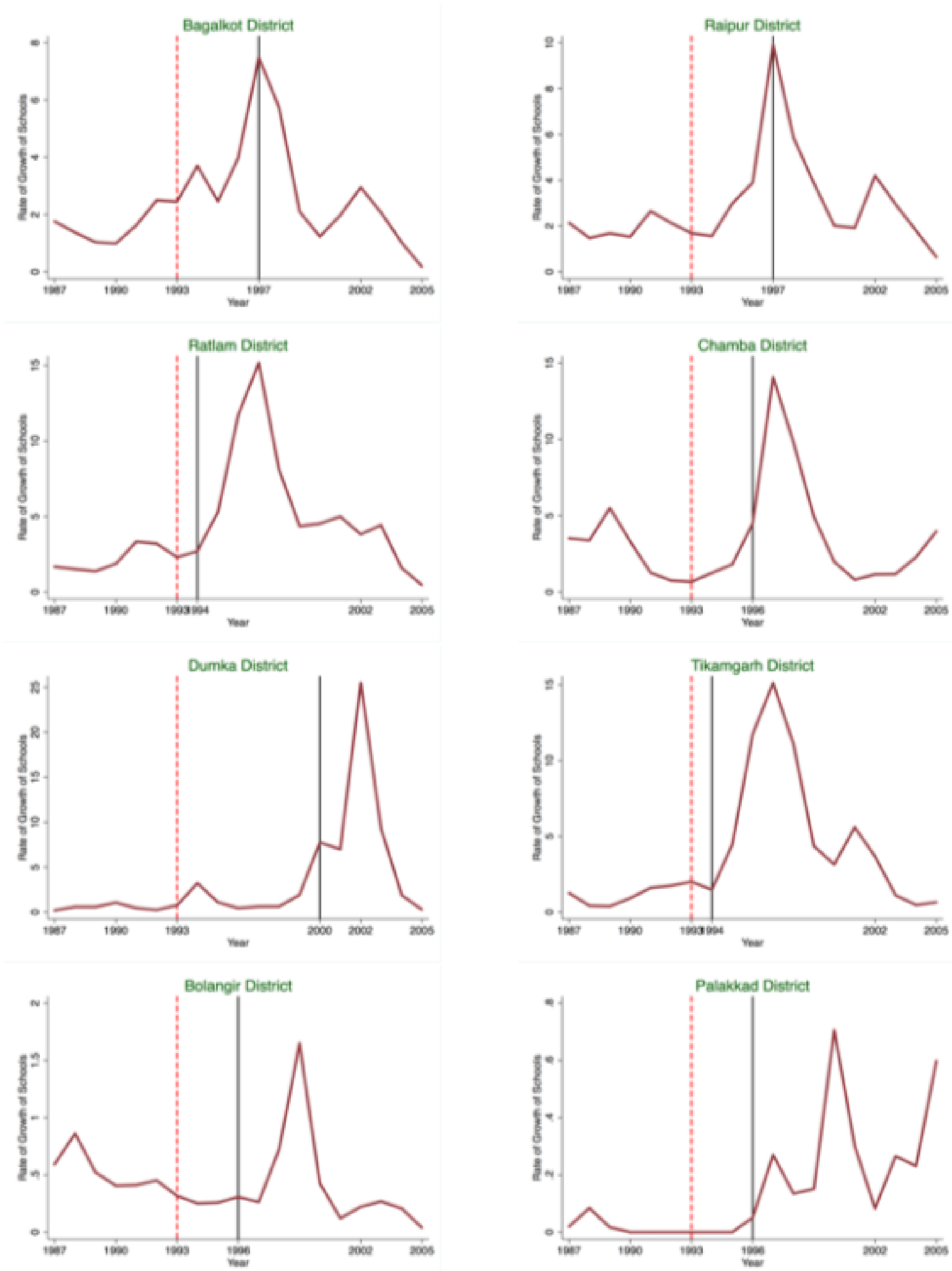

Figure 1.3: Yearly rate of growth of school construction plotted against time – Treatment Districts

The graphs in this figure illustrate that the peak in school construction growth in treatment districts is better *predicted* by the year of programme implementation that I infer using the government archival data, rather than the uniform start year of 1993-94.

Data : DISE 2005.

LINK TO RESULTS SECTION
LINK TO EMPIRICAL STRATEGY SECTION

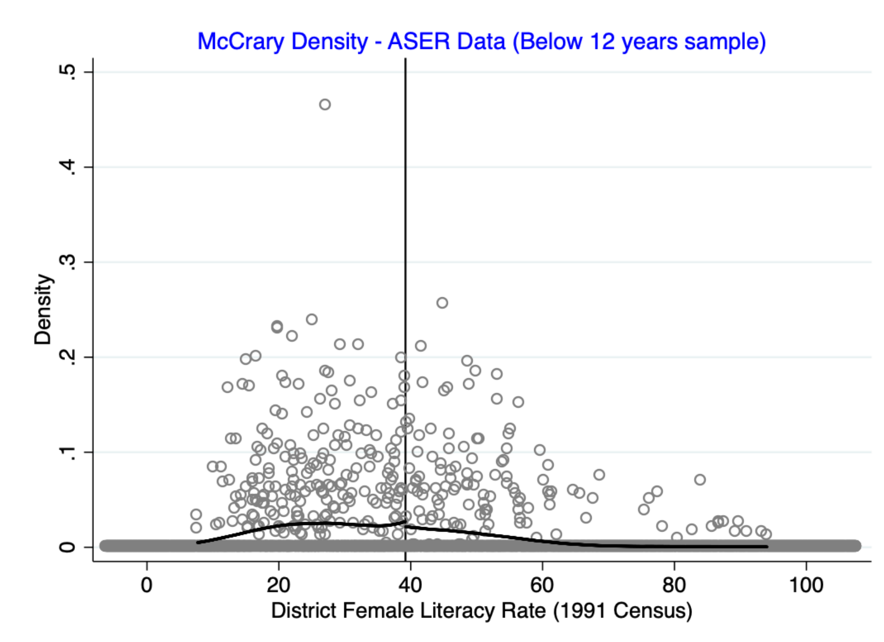

Figure 1.4: McCrary Density test using ASER data.
(described in the data section) as per McCrary, 2008. The associated test statistic is 0.003, and the p-value is 0.42. LINK TO RD VAIDITY SECTION

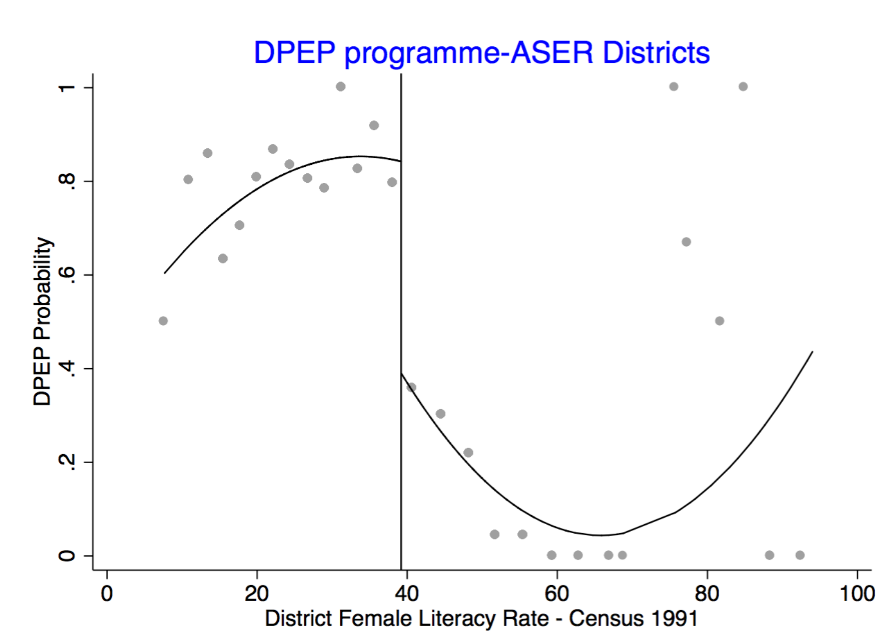

Figure 1.5: Probability of Receiving DPEP Programme.
The graph shows the discontinuity of treatment assignment at the cutoff of 39.2 percent in terms of District Female Literacy Rate. Data Source : ASER data combined with information in government archives. LINK TO RESULTS SECTION
LINK TO EMPIRICAL STRATEGY SECTION

Figure 1.6: Impact on Mathematics test scores.
This is a graphical representation of the estimate presented in table 1.4. Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. LINK TO RESULTS SECTION

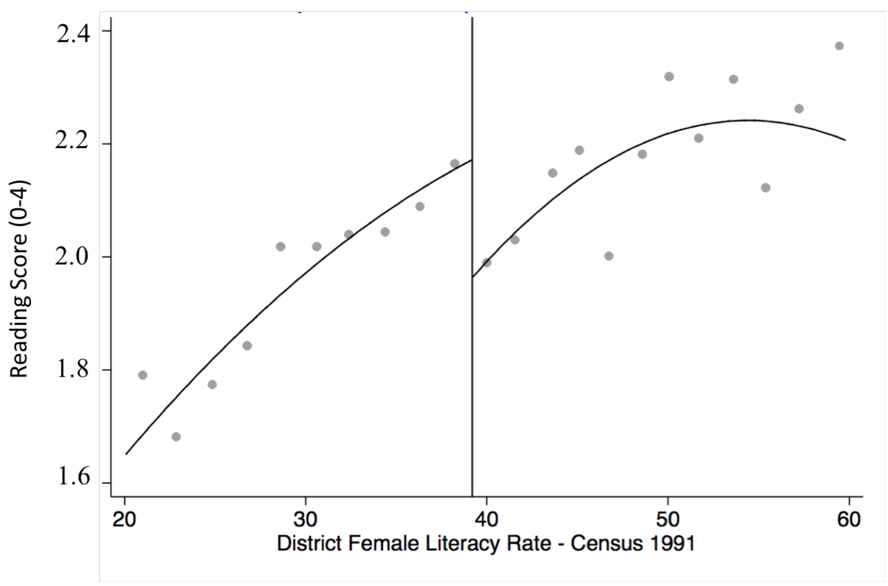

Figure 1.7: Impact on Reading test scores.
This is a graphical representation of the estimate presented in table 1.4. Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. LINK TO RESULTS SECTION

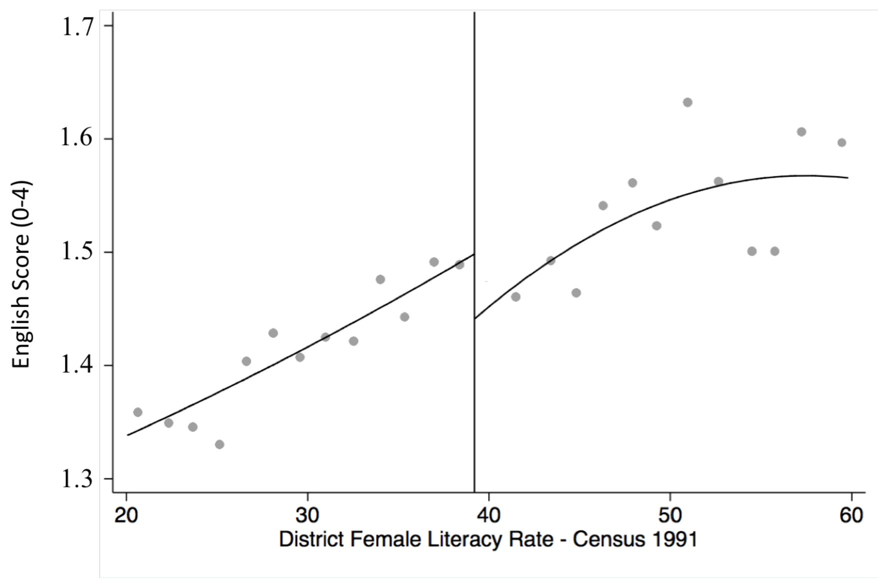

Figure 1.8: Impact on English test scores.
This is a graphical representation of the estimate presented in table 1.4. Based on author's calculations using the Annual Survey of Education Report (ASER) individual level data for the years 2007–2014. The RD point estimates are constructed using the triangular kernel, local polynomial of order 2 and with one common CER–optimal bandwidth selector. All specifications control for the age of the child, ages of both parents, rainfall shocks in–utero/birth year of the child and dummy variables for state and year of data collection. The standard errors are robust and are clustered at the district level. LINK TO RESULTS SECTION

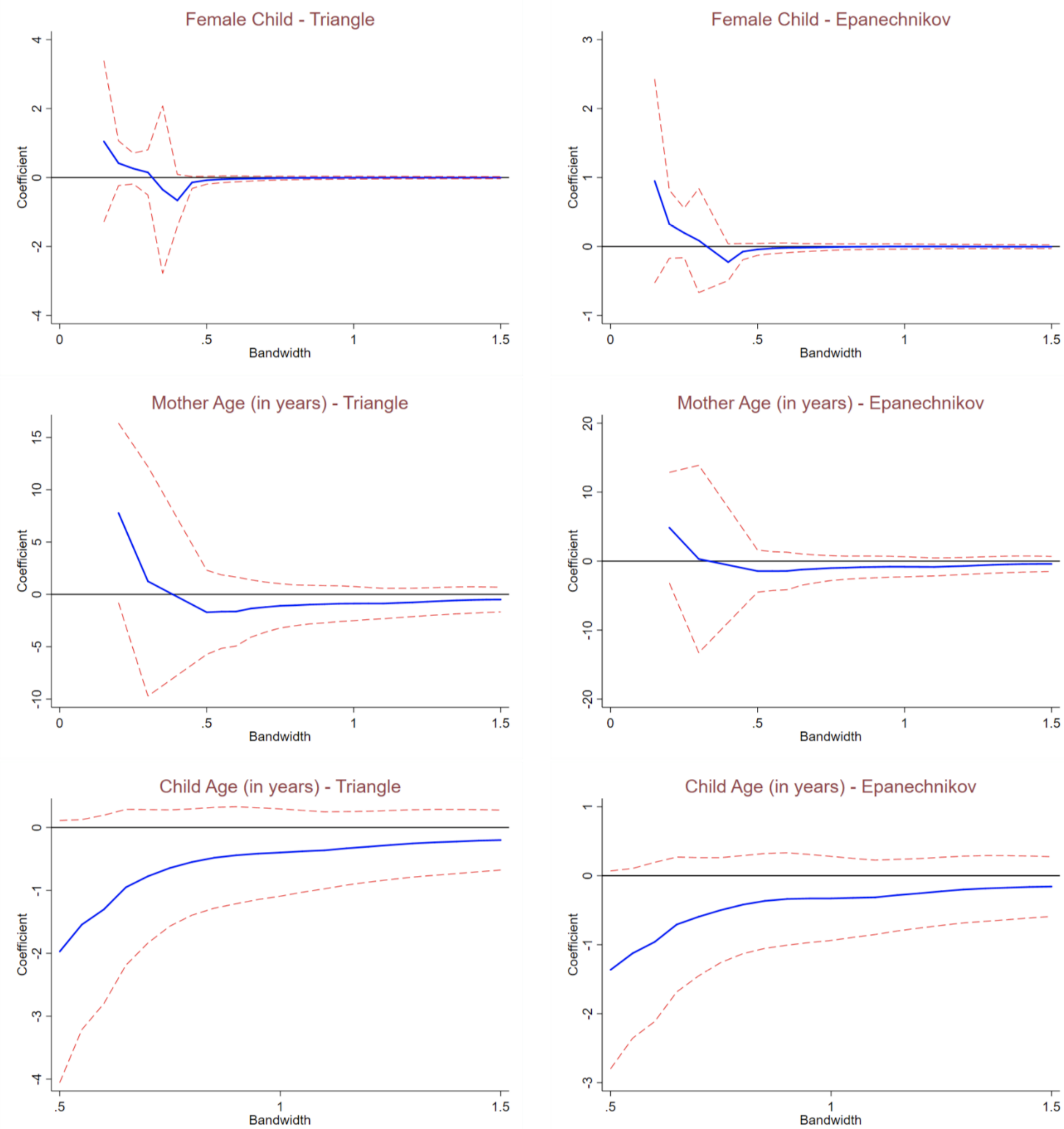

Figure 1.9: Discontinuity in Pre-Determined Outcomes

These graphs show that the DPEP programme had a statistically insignificant (i.e. indistinguishable from zero) impact on these outcomes. The staring year for the treatment districts comes from the government archives, while that for the control districts comes from the average starting year of treatment districts within the same state. The left panel shows graphs of the RD impact estimate for an outcome using **triangular** kernel with a **polynomial of degree 2**. The right panel does the same with an epanechnikov kernel. In both graphs the coefficients are estimated at different bandwidths, where bandwidths are increased in steps of 0.05. All estimates use **robust–bias standard errors** with **clustering at the district level**. LINK TO RESULTS SECTION

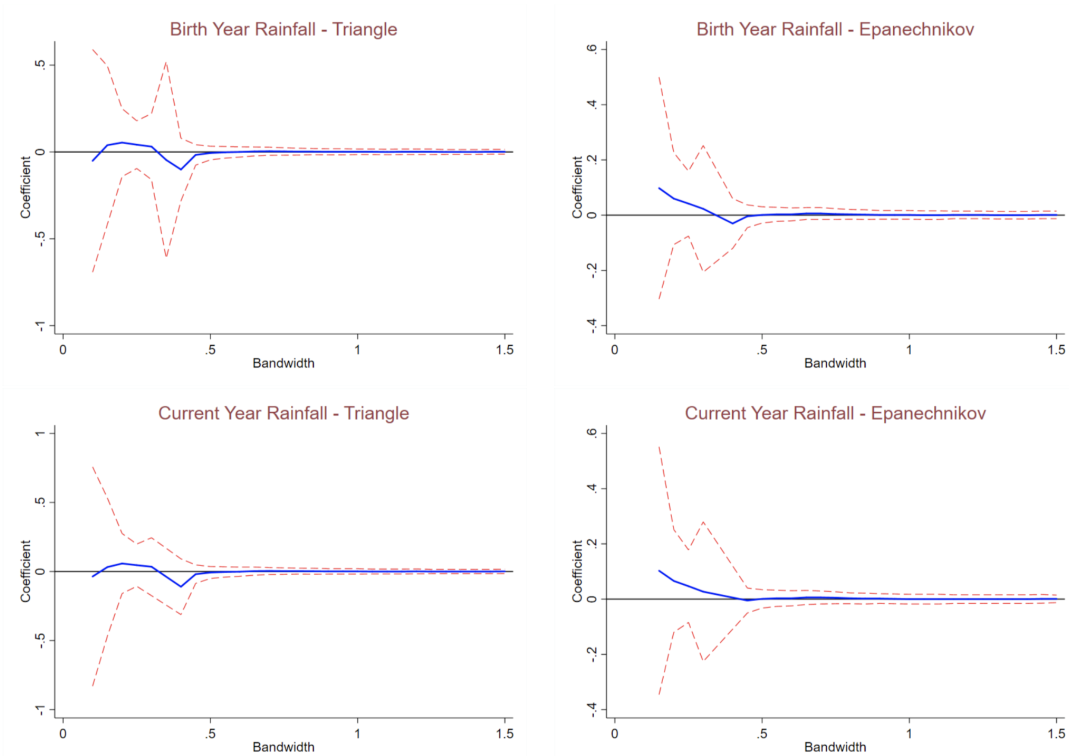

Figure 1.10: Discontinuity in Pre-Determined Outcomes
These graphs show that the DPEP programme had a statistically insignificant (i.e. indistinguishable from zero) impact on these outcomes. The staring year for the treatment districts comes from the government archives, while that for the control districts comes from the average starting year of treatment districts within the same state. The left panel shows graphs of the RD impact estimate for an outcome using **triangular** kernel with a **polynomial of degree 2**. The right panel does the same with an epanechnikov kernel. In both graphs the coefficients are estimated at different bandwidths, where bandwidths are increased in steps of 0.05. All estimates use **robust–bias standard errors** with **clustering at the district level**. LINK TO RESULTS SECTION

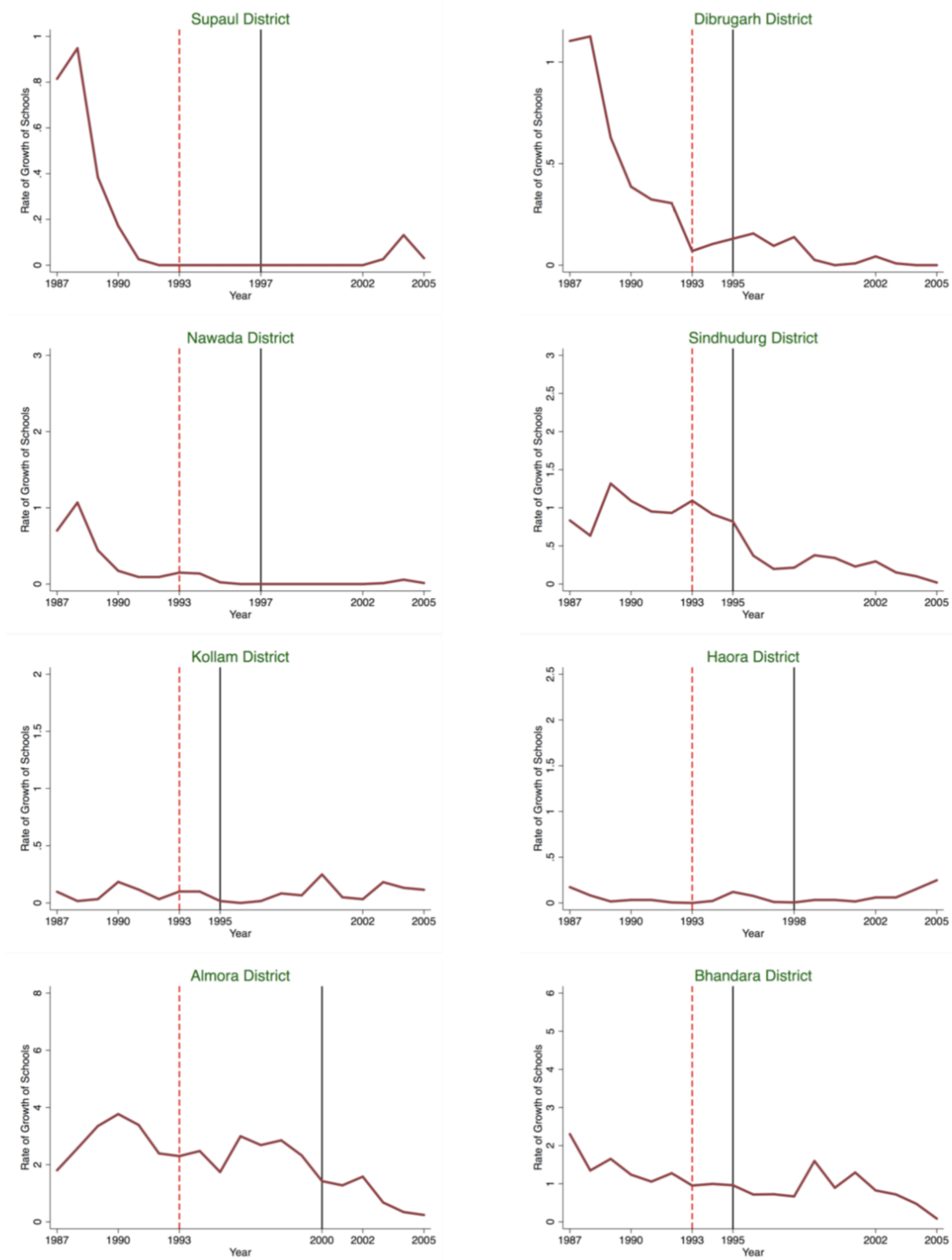

Figure 1.11: Yearly rate of growth of school construction plotted against time – Control Districts

The graphs in this figure illustrate that there was no upward trend in school construction in the control districts around the time the DPEP programme was implemented. Data : DISE 2005.

LINK TO RESULTS SECTION

CHAPTER 2

WHAT YOU LEARNED BY SECOND GRADE MATTERS: A COMPARATIVE STUDY ON HUMAN CAPITAL FORMATION IN MADAGASCAR AND SENEGAL

(with Heidi Kaila & David Sahn)

2.1 Introduction

Cognitive skills have important implications not only for individual wellbeing (Heckman, 2006), but also for economic growth (Hanushek and Woessmann, 2008, Hanushek, 2013). In this paper, we study the determinants of grade attainment and academic performance among young adults in two francophone African countries, Madagascar and Senegal. We examine the importance of second–grade skills, measured by test scores on math and French, in determining the educational level and test scores of young adults[1]. Additionally, we explore the role of other conditions, such as health, wealth, and parental education in shaping the outcomes. We do so by using two unique and comparable panel surveys from Madagascar and Senegal that follow children from second grade until they are young adults. The cohorts are followed over a period of 15 years in Madagascar and 17 years in Senegal, an unusually long period for survey data, especially in the African context.

Our study adds to the literature in economics and psychology that discusses the positive relationship between childrens performance in elementary school

[1]We use the term skills to refer to academic skills, that is French and math skills, measured in second grade and early adulthood. Even though these skills strictly measure academic performance, we consider them a proxy for cognitive skills formation in general.

and their academic performance later in life (Cunningham and Stanovich, 1997, Feinstein, 2003, Bourne et al., 2007, Duncan et al., 2007[2]). For the most part, these studies measure academic performance in terms of school attainment and test scores. We follow this literature and compare the relative importance of the impact of early–life math and language scores on school attainment and young adult test scores across two African countries. Our results closely mirror those of Duncan et al., 2007, who used data from the United States, Canada, and the United Kingdom to show that math skills at school entry were a stronger predictor of later achievement than language skills[3].

Apart from contributing to the literature on the persistence of academic skills from school entry to young adulthood, we also examine how second–grade skills affect school progression. Glick and Sahn, 2010 found that skills in early primary school (second grade) in 1995–6 were strongly positively associated with school progression (measured through grade repetition) eight years later. Similarly, Singh, 2017, using panel data from India found that skills in primary school increased the likelihood of completing secondary schooling. In our analysis, we document similar relationships with the added dimension of using data sets that span a longer period of time, as compared to the aforementioned studies. This allows us to examine the impact that second grade test scores have on human capital formation of a cohort of young adults who were last surveyed in their early twenties, which previous studies have not been able to do.

While the primary focus of our research is on the relationship between

[2]For a review article on effects of early life attributes on adult outcomes, see Heckman and Mosso, 2014.

[3]Our results should be interpreted while keeping in mind that French, which is the language of instruction in our sample, might not be the first language learned by the children at home. This contrasts with the case of Duncan et al., 2007, who presented evidence from predominantly anglophone developed countries where the language of instruction was likely to have been the same as the primary language at home.

second–grade skills and human capital outcomes later in life, we are also interested in the role played by health in determining later–life outcomes. Therefore, we include height in our models to examine the role of childhood health in shaping adult outcomes. The additional benefit of including height in our models is that it controls for the confounding effect it could have on our primary question of interest: the relationship between skills in second grade on grade attainment and academic skills in young adulthood. In this regard, we build upon the related evidence of the impact of health, measured using adult height, on human capital formation. Adult height is determined largely by a confluence of genetic and environmental factors, especially in–utero and during the first 24 months of life (Tanner, 1979, Strauss, 1997, Currie and Vogl, 2013). It has been found to be strongly associated with a variety of adult socioeconomic outcomes (Case and Paxson, 2008a, Lundborg et al., 2014, Vogl, 2014, Sohn, 2015, LaFave and Thomas, 2017). Case and Paxson, 2008a found a strong relationship between adult height and earnings, with the pathway being higher cognitive skills of taller individuals, leading them to select into occupations with higher earnings. Lundborg et al., 2014 also found that height at age 18 has a significant impact on wages, even after controlling for cognitive and noncognitive skills in Sweden. Similarly, Vogl, 2014 and Sohn, 2015 found a large height premium in wages in Mexico and Indonesia, respectively[4].

In our work we also control for family background, particularly wealth and parental education, as they play a critical role in the process of mental and physical development. In doing so, our paper builds on the literature that provides

[4]With respect to childhood height, it has not only been shown to be a strong determinant of adult height (Case and Paxson, 2008b), but Alderman et al., 2006 reported that childhood health status (as measured by height–for–age) had a positive impact on completed schooling as an adult. Behrman et al., 2014 also found that height–for–age at age six affected adult cognitive skills, and further made the point that excluding height from a cognitive production function will result in the overestimation of the effect of schooling on cognitive ability.

strong evidence of parents education affecting grade progression and academic skills of their children, both in developing and developed countries (Cunha and Heckman, 2007, Todd and Wolpin, 2007, Cunha et al., 2010, Glick et al., 2011, Behrman et al., 2014, Jones et al., 2014, Marchetta and Sahn, 2016, Behrman et al., 2017). In the context of India, Helmers and Patnam, 2011 found that parental investment had an impact on skill levels of children of primary school age or lower. Glick et al., 2011 and Jones et al., 2014 used data from sub-Saharan African countries to find that parental background plays a significant role in the determination of childrens ability, and similarly, children of educated (and nonpoor parents) have been found to perform much better than their peers on cognitive tests (Dumas and Lambert, 2010).

We also control for wealth by having an asset index in our models – there is significant evidence that material wellbeing matters for education and cognition outcomes (Cunha and Heckman, 2007, 2008, Todd and Wolpin, 2007, Helmers and Patnam, 2011, Schady et al., 2015). However, in most studies from developing countries, income, expenditure, or wealth is measured contemporaneously with the outcome of interest. This association between contemporaneous household resources and skills needs to be interpreted with caution, since causality runs in both directions. In contrast to the preponderance of the related literature, we explore the relationship between wealth (measured by assets) at the time children were in second grade on their human capital as young adults. Although a causal interpretation is still not possible, especially due to the impact of unobserved heterogeneity, using wealth lagged 15 years eliminates, at least, the concern of reverse causality in our context.

To explore the importance of the aforementioned socioeconomic factors, height, and skills in second grade on young adult outcomes, we build upon a standard human capital production function framework (Todd and Wolpin, 2003, 2007, Cunha and Heckman, 2007, Cunha et al., 2010), in which skills are developed over time and are a function of inputs received by the child, such as parents education, wealth, and the schooling environment. The unique nature of our dataset permits us to assess these relationships while controlling for a variety of household – and school–level factors. We also discuss heterogeneous effects on dimensions related to gender, height, and household wealth.

Another distinguishing characteristic of our work is that we study these research questions in the context of two sub-Saharan African countries using comparable data sets. There is a relative lack of evidence of the type that we present here with respect to Africa, mostly due to the lack of availability of long–term panel data sets. It is even more lacking with respect to being able to provide comparative evidence across countries. Our paper bridges this gap in the literature.

To strengthen the comparability of our results from Madagascar and Senegal, we use Item Response Theory (IRT)[5] to create joint test score indices for the two countries. We use these scores to compare performance across the two countries at two points of time (childhood and adult life). Such comparisons of human capital formation, especially from developing countries, are quite rare (Jones et al., 2014, Schady et al., 2015, Singh, 2017). Our analysis is different from these papers insofar as our data covers a longer time period, effectively spanning the entire course of the schooling experience from second grade to early adulthood.

We find that human capital in young adulthood is strongly associated with

[5]Appendix A explains the construction of the IRT scores and how this facilitates comparison of test scores across the two countries.

skills in second grade in both countries. We observe heterogeneous effects of math and French skills in second grade, where math scores have a stronger relationship with later–life outcomes. These results confirm patterns observed in Duncan et al., 2007. The results also suggest that taller individuals have higher test scores, evidence similar to that found in Case and Paxson, 2008a and LaFave and Thomas, 2017. We also find that height plays an important role in grade attainment among young adults in Senegal, but the coefficient estimate is not statistically significant in the Malagasy case. This coefficient of child health is obtained after controlling for skills in second grade. In addition, our results indicate that wealth of the household when children are in second grade is associated with schooling and skills measured more than 15 years later. Furthermore, we find that parents education matters more in Madagascar than in Senegal. Finally, our heterogeneity analysis shows that second grade test scores are more strongly associated with later–life outcomes for shorter individuals and females, groups that are potentially more vulnerable (akin to the analysis of Glewwe et al., 2017). This implies that poor performance in second grade is more detrimental to outcomes as an adult for certain groups, as compared to others.

The remainder of this paper is organized as follows. Section 2 presents the country contexts, while Section 3 expands on the data and some comparative descriptive statistics on Madagascar and Senegal. In Section 4 we discuss the conceptual framework and the empirical strategy used. In Section 5 we present the results and discuss robustness checks in Section 6. Finally, in Section 7 we draw conclusions.

2.2 Context

Our comparative study examines the production of human capital among young adults in two poor sub–Saharan African countries, Madagascar and Senegal. Although these countries differ along many dimensions, they share many similarities. Both are low–income countries that struggle with low school attainment and primary school completion rates. This is still the case despite significant improvements in primary school completion rates over the study period (1996–2012), from 40 to 59 percent in Senegal and from 31 to 70 percent in Madagascar. In the same period, gross enrollment rates in primary schools have also risen, in Senegal from 59 percent to 81 percent, and in Madagascar from 86 percent to 145 percent (WorldBank, 2016). Additionally, in the 1990s, grade repetition and dropout rates were high in both countries (Michaelowa, 2001, Glick and Sahn, 2010). The educational systems in these countries are modeled after the French system, and the primary language of instruction is French.

Although children in Madagascar and Senegal are exposed to potentially similar schooling systems, they differ critically in the opportunities they may encounter and the extent to which their background matters for their achievements in later life. Madagascar is an island economy that has experienced almost two decades of political turmoil, with average GDP per capita growth being zero during the period of our study (World Bank 2016). In contrast, Senegal is one of the more dynamic economies in West Africa, with GDP per capita growth averaging 1.2 percent from 1995 to 2012. Likewise, the poverty headcount ratio has increased slightly in Madagascar, to nearly 75 percent in 2010. However, in Senegal the headcount ratio stood at 47 percent in 2010 (WorldBank, 2016). Madagascar has lower levels of intergenerational mobility of education and occupation than does Senegal, as well as most other African countries (Bossuroy and Cogneau, 2013, Azomahou and Yitbarek, 2016). Further, Glick et al., 2011 found that parents education and schooling are important determinants of learning in Madagascar. In Senegal, Glick and Sahn, 2009 showed that, conditional on a childs level of schooling at 14 to 17 years of age, having better educated parents or a higher level of household resources have only modest benefits for academic performance. They found similar results for school–level variables.

2.3 Data

The first round of the survey was conducted in 1995–6 in Senegal and in 1997–8 in Madagascar. Math and French tests were administered to children at the beginning and end of second grade, when the children were between 7 and 10 years of age[6]. These school–based tests were administered as part of a multicountry study called the Program on the Analysis of the Conference of Francophone Ministers of Education, which is referred to by its French acronym, PASEC[7]. Both urban and rural communities were included in the PASEC school–based sample, which was designed to be a nationally representative selection of schools from communities throughout the country. This involved randomly selecting communities from throughout the country. In cases where there were multiple schools in the selected community, one school was chosen at random to be part of the sample. Despite the intention to draw a

[6]Some children were older or younger because of early or delayed enrollment.

[7]In French, the study name is Programme danalyse des systmes ducatifs de la Confemen. They were conducted under the authority of the Conference of Education Ministers for Francophone Africa, CONFEMEN. For more information on the PASEC, see PASEC (2016) and Michaelowa, 2001.

representative sample of students from the entire country, there are two important qualifications. First, while the communities, and then schools within the communities were randomly selected, the PASEC tests were school–based tests, which meant they were were administered only to school–going children. As with all school-based testing for such national testing programs, as well as large cross–national surveys such as the Program for International Student Assessment (PISA)[8], the sample is restricted to those who were enrolled in school. Hence, the sample is not representative of the entire cohort of children in the relevant age range, since there were some children in each country who had never been enrolled in school and other children who had left or dropped out of school before second grade. Second, schools were selected for the sample only if they had a class size of at least 20 students, although that could include multi–grade classrooms. In practice, almost all schools in Senegal had sufficient number of children in the classroom and thus were eligible to be included in the sample, but, in Madagascar, we found that the smallest and most remote communities in the country were underrepresented in the sample of schools that was selected in the mid–1990s.

A subset of the children in the PASEC surveys in Madagascar in 1997–98 and Senegal in 1995–96 were followed up in 2012–13. The 2012–13 data sets are referred to as the Life Course Transition of Young Adults Surveys. The young adults were, on average, 22 years old in Madagascar and 24 years old in Senegal at the time of the surveys in 2012–13. The children in this long–term cohort were randomly selected from slightly less than half the original clusters included in the PASEC surveys of the mid–1990s[9]. Our final sample that spans the period of over 15 years includes 333 and 447 children who were in second grade in the

[8]http://www.oecd.org/pisa/test/

[9]Selecting of subset of communities was necessitated by budgetary constraints at the time.

1990s in Madagascar and Senegal, respectively[10].

As indicated above, skills assessments, in the form of math and French tests, were administered in both survey rounds. It should be noted that the tests administered in the two countries were either the same or had a subset of common questions (more details in B.1). However, the tests for children and for adults were different, reflecting the different periods in the cohort members life courses. The presence of common questions in the tests administered in the two countries allows us to construct test scores based on the Item Response Theory (IRT), using the joint distribution of the two test scores. The parameters of IRT are estimated jointly for the common items, which renders scores comparable across the countries for a given time period[11][12]. This enables us to conduct a descriptive comparison of the test performance across the two countries. Additionally, the IRT score also has the benefit of being a cardinal measure of test performance, as opposed to the more commonly used measure of percentage of correct answers, which is merely an ordinal measure.

Figures 2.1, 2.2 and 2.3 plot the cumulative density functions (CDF) of the test scores for the two time periods, using the IRT estimates from the joint distribution of the test scores. The distribution of second grade composite scores (Figure 2.1) for Madagascar first order stochastically dominates the distribution for Senegal. This pattern holds for both math (Figure 2.2) and French (Figure 2.3) scores separately. By 2012, there had been considerable convergence in the

[10]The main reason for the smaller sample size in Madagascar is that fewer communities with PASEC schools were followed up in Madagascar than in Senegal. More information on attrition, along with robustness checks for attrition, is provided in Appendix C.

[11]The details of which tests were merely similar, and which were the same, are given in Appendix A along with the description of the IRT methodology.

[12]While we can compare the performance across countries, we cannot do so across time. This is because as the tests administered to adults have no common items with the tests administered to children.

distribution of the scores across the countries. In Figure 2.4 we provide descriptive evidence on the relationship between the second grade and early adulthood scores, using the jointly estimated (comparable) IRT scores. A clear implication from Figure 2.4 is that the relationship between the second grade and early adulthood scores is stronger in Senegal than in Madagascar, as the slopes of the curves are steeper. Furthermore, the narrower confidence bounds further illustrate the strength of this relationship in Senegal, compared to Madagascar[13]. Figures 2.5 and 2.6 show the nonparametric relationship of test scores measured for young adults in relation to height in young adulthood. We can see that, in both countries, the test scores are increasing in height, implying taller individuals did better in cognitive tests as adults. This is similar to the findings of Case and Paxson, 2008a and LaFave and Thomas, 2017.

While the IRT scores based on the joint distribution are especially useful for descriptive comparisons of test scores across countries, in the regression analysis presented below, we employ IRT scores that were estimated separately for each country. This allows us to better model changes in test scores within country and across time, since the country–specific IRT scores are a better estimate of the country–specific measures of ability.

Tables A2.1 present summary statistics of the variables of interest in Senegal and Madagascar, respectively[14]. The test score variables are the country–specific IRT transformations with means close to zero, both in the 1990s and in 2012. In Senegal the adult sample has completed an average of 9 grades in school, com-

[13]The larger confidence bounds in the low and high ends of the test score distribution reflect the fact that there are less observations at the ends of the distributions.

[14]The sample sizes vary slightly depending on the dependent variable, due to missing observations in the test scores. The highest grade completed in Madagascar (Senegal) is available for 333 (447) individuals; the 2012 Math scores are available for 318 (447); the 2012 French scores for 312 (381) individuals; and the joint math and French test scores are available for 310 (381) individuals in Madagascar.

pared to 10 grades in Madagascar. In Senegal the sample has a slight majority of males, whereas in Madagascar, it is the opposite. On average, the Senegalese sample is 24 years of age, slightly older than the Malagasy sample of 22 years of age. This is consistent with the second–grade baseline data having been collected two years earlier in Senegal. In addition, the Malagasy sample is roughly 10 cm shorter than the Senegalese sample. The discrepancy is similar to that found in the DHS data (Subramanian et al., 2011). The difference is large and is likely to be partially explained by different ethnic compositions of the populations, where the largest ethnic groups in Madagascar are of Asian descent. Additional information on household characteristics were collected in the original PASEC surveys conducted in the mid–1990s, including a detailed listing of all the assets owned by the household. This allows us to create a household asset index using factor analysis.

One concern with the data sets we use is attrition. We faced the challenge of returning to communities many years after the original PASEC survey was conducted in the mid–1990s and searching for the original children over 15 years later. The difficultly of doing so was exacerbated by the exceedingly low living standards, volatility in economic conditions, and the constant social transformation in the original communities surveyed. Despite these challenges, we began by randomly selecting a subset of the communities and a subset of the children in each community to be included in our follow up surveys. Despite our efforts to identify as many of the children as possible, the attrition rates in both in Senegal and in Madagascar were just under 50 percent for the 17– and the 15–year intervals, respectively.

To better understand the implications of this attrition rate, we compared the sample of children in the cohort included in our analysis with those in the original PASEC samples, recalling that the original surveys conducted in the mid-1990s were designed to be representative of school–aged children in the countries. Appendix Table A2.2 show a comparison of means for key variables between these samples. For Senegal, the children in the panel had slightly lower second–grade test scores on average than the children not in the panel, and they also came from households with slightly lower asset scores. For Madagascar, the children interviewed in second grade, and as adults, came from households with slightly less wealth. These individuals are also slightly younger than the overall sample. In Madagascar, however, we see no systematic differences in test scores. We find that the differences arise from the fact that an attempt was not made to reach all communities in the follow–up, rather than arising from attrition within a community (Tables A2.3 and A2.4). To alleviate concerns related to attrition and sample selection, we conducted a robustness check of our results by estimating inverse probability weighted regressions, where the weights are calculated based on the estimated probability of being in a community that was included in our follow–up surveys. This check is based on different characteristics in the baseline data. This is discussed in detail in Section 6. The results from this exercise show that the results are robust to the attrition and sample selection-related adjustment.

Despite the checks intended to alleviate concerns over attrition and sample selection for follow–up, we want to emphasize that we do not make any claims regarding our cohort being representative of the entire population in the age group of the cohorts in the two countries. This is because, as noted above, this is a school–based sample where children in second grade were administered these tests. Therefore, by design, the sample excludes children who had not completed at least one year of schooling at the time of the first survey and those from the smallest, most remote villages in Madagascar. We note, however, that we have a unique, long–term panel data set from two developing sub–Saharan African countries where we were able to follow individuals cognitive ability, as measured by test scores, from second grade until young adulthood and explain the evolution of scores with information on socioeconomic background. Therefore, despite the recognized limitations in terms of sample size and attrition, there is much to learn from these data sets[15].

2.4 Conceptual framework

In this section, we first present a simple theoretical framework of the cognitive production function, and then, present the empirical framework that we use to estimate it.

2.4.1 Theoretical Framework

Our theoretical framework builds on the work of Todd and Wolpin, 2003, 2007, which is also the analytical point of departure of Aubery and Sahn, 2014, Fiorini and Keane, 2014, Singh, 2017. We also draw on literature, studying the relationship between height and later–life outcomes, which for the most part, finds a strong association between height and cognitive skills in adulthood (Case and Paxson, 2008b, LaFave and Thomas, 2017).

[15]We should note that other panels of this type of duration and detail from developing countries, such as the Guatemalan studies (Grajeda et al., 2005, Behrman et al., 2014) and Young Lives studies (see link), suffer from similar attrition problems. For a discussion, see Alderman et al., 2001

Considering childhood and adulthood as two periods of life, the following would be a simple illustration of the two–period mechanism pertaining to grade attainment

$$Y_2 = f\left(\beta_1 A_1\left(\mu_0\right) + \beta_2 P_1 + \beta_3 H_2\left(n_1, \mu_o\right) + \beta_4 S_1\right) \tag{2.1}$$

where the grade attainment Y_2 in Period 2 is a function of cognitive ability A_1, and height H_2 in Period 2, which is largely determined by cumulative health/nutritional endowments and socioeconomic factors (particularly, in the first few years of life) denoted by n_1 (Martorell and Habicht, 1986b). Also, both ability and height are functions of a genetic component, $_o$, at the time of conception. In Period 1, S_1 denotes the school inputs and P_1 denotes parental investments, including factors such as household wealth and the education of parents.

If expressed in the form of skill accumulation, the model is

$$A_2 = g\left(\gamma_1 A_1\left(\mu_0\right) + \gamma_2 P_1 + \gamma_3 H_2\left(n_1, \mu_o\right) + \gamma_4 S_1\right) \tag{2.2}$$

2.4.2 Empirical Framework

The simplest empirical counterpart of equation 2.1 is a reduced form model, which can be estimated using an OLS model:

$$Y_{i,2012} = \beta_o + \beta_1 A_{i,1996} + \beta_2 Height_{i,2012} + \beta_3 HH_i + \beta_5 X_i + \gamma_j + \varepsilon_i \tag{2.3}$$

In this regression, $Y_{i,2012}$ stands for the highest grade attained by the cohort member in 2012, and $A_{i,1996}$ stands for a measure of childhood skills, which in our case is measured using math and French scores at the beginning and end of second grade (called pretest and posttest, respectively). The time index used in equation 2.3 is 1996, which is the year corresponding to the Senegal data, while in the

Madagascar data it is 1998. $Height_{i,2012}$ refers to height measured in early adulthood in 2012, which is a function of both health inputs received over the life course, particularly, in utero and in early childhood, as well as genetics (Martorell and Habicht, 1986b). We include it in our model, as evidence suggests that adult height is strongly related to outcomes in adulthood; thus, omitting height from the model could potentially lead to inflated coefficient estimates for other covariates in the model (Case and Paxson, 2008b, LaFave and Thomas, 2017). HH_i is a vector of household–level (time–invariant) inputs; γ_j are school fixed effects, corresponding to school j; and X_i denotes time–invariant control variables.

Estimating equation 2.2 leads to a very similar reduced form regression:

$$A_{i,2012} = \beta_o + \beta_1 A_{i,1996} + \beta_2 Height_{i,2012} + \beta_3 HH_i + \beta_4 X_i + \gamma_j + \varepsilon_i \tag{2.4}$$

Our dependent variables $A_{i,2012}$ are performances on French and math tests in 2012. We model these test score outcomes individually, as well as a composite score.

This setup is analogous to a value–added (VA) specification in which current test scores are regressed on earlier period outcomes and other determinants. Although value–added models have primarily been used to analyze skill acquisition from one grade to another, often focusing on estimating teacher and school characteristics and quality of learning, our paper differs in an important way: we are interested in explaining skills in early adulthooda time during which the cohort is no longer in school using skills from a previous period (second grade)[16]. Since our empirical specification includes lagged inputs and lagged

[16]See Fiorini and Keane, 2014 for an overview of different specifications of VA models to

achievement, it can therefore be thought of as a combination of the cumulative and value–added models, as described in Fiorini and Keane, 2014. It generalizes the value–added (VA) model, which was preferred in Todd and Wolpin, 2007, because it minimized the out–of–sample root mean squared error[17]. Although our framework is statistically equivalent to a VA model with lagged inputs, it differs conceptually, as the time period between our waves is fairly large, 15–17 years.

Our model takes school inputs into account through the inclusion of school fixed effects. Each individual is assigned to the school that he/she attended in second grade. The school fixed effects control for all time–invariant, school–level factors, as well as class-level, time–invariant unobservables, as our data has only one class per school. The fixed effects also control for time–invariant, community–level factors, due to the one–to–one correspondence between schools and communities. Consequently, our empirical specification compares children who attended the same school (and class) when in second grade, after controlling for other household and individual covariates. Hence, it is likely that the children were exposed to roughly the same socioeconomic and environmental factors in their childhood, thus making the comparisons even more relevant. We include in the models the asset index of the household where the young adult lived when in the second grade. The coefficient can then be interpreted the effect of wealth in early childhood on later–life outcomes. We also use height in early adulthood, a proxy for the lifetime cumulative health status (but predominantly influenced by the period from conception to 24 months of

explain cognitive skill formation for school–aged children with contemporaneous and lagged inputs.

[17]In addition, Fiorini and Keane, 2014 also discussed data intensiveness of these procedures and the associated sample size issues in their analysis. We also face similar challenges but still end with nearly the same sample size as their specifications.

age) on young adult test scores. Additionally, we control for some individual–specific attributes, such as parents education, which further reduces concerns regarding omitted variables in this specification.

We measure childhood academic ability of the individual using second–grade math and French test scores. We obtain information on scores of pretests that were administered to students at the start of second grade (1995 in Senegal and 1997 in Madagascar), and posttests that were administered at the end of second grade (1996 in Senegal and 1998 in Madagascar). In addition to capturing the role of cognitive ability and genetic factors that contribute to test performance in the second grade, these early–life test scores are a function of household and school inputs that the children received from the time of conception until the second–grade test was conducted. Although we have a choice of using scores from tests conducted at the start and end of the school year, we use the latter in our specifications, because the tests are comparable across the two countries. We also show models with composite scores (using both the pretests and posttests) during second grade as a robustness check (2.9).

We also estimate a model with only lagged inputs, as well as a model that excludes lagged test scores, which in Todd and Wolpin, 2007 and Fiorini and Keane, 2014 is referred to as the cumulative model. This model assumes that the lagged inputs incorporate the innate ability and unobserved inputs. Our estimations clearly show that this is not the preferred specification, as the lagged test scores are statistically significant[18].

[18]Another potential specification for studying the effect of lagged inputs on skills in early adulthood is to employ a fixed effects framework, as in Fiorini and Keane, 2014. The underlying assumption is that the lagged coefficient of the test score is equal to one (Singh, 2017). Our results show that this is not a valid assumption, as the coefficient estimates are much lower, as in Singh, 2017 and Fiorini and Keane, 2014; but more important, this approach is not feasible due to the fact that we do not have time–varying inputs in our regressions.

In the main specifications, we use test score variables that are created based on Item Response Theory (IRT). We use this method to create three separate sets of test scores for each round of survey in each countrymath, French, and composite scores. In latter specifications, we explore whether math and French test scores obtained during the second grade are equally strong predictors for adult skills, or if, as found in some literature from the predominantly English–speaking world, math ability is a stronger predictor of skills in later life (Duncan et al., 2007, 2011).

2.4.3 Correcting for Measurement Error

We can only control for the observed individual, household, and school factors; thus, unobserved factors are part of the error term. These unobserved factors might in turn be correlated with both our outcome of interest (such as later–life schooling) and the childhood test scores, thus leading to endogeneity bias. It is important to note that the previously discussed instrumental variable strategy does not necessarily correct for this endogeneity bias and simply addresses systematic measurement errors. Thus, similar to other papers that have looked at childhood ability and how it affects outcomes in adult life, we rely on an important set of controls to at least mitigate endogeneity concerns. While we acknowledge the possibility of endogeneity in our specification, we also note that several recent papers, which have compared value–added estimates of the type explored here and estimates from experimental or quasi–experimental analyses, have mostly concluded that the non–experimental estimates are unbiased, when compared with estimates from experiments (Angrist et al., 2013, Kane et al., 2013, Deming et al., 2014, Deming, 2014). Also, the long duration of our panel mitigates concerns related to endogeneity and, to our knowledge, there is no research that covers this large a span of time that has fully addressed these endogeneity concerns, and which could only be done if there was some sort of experimental design implemented during the original time periodin our case, the mid 1990s. To further assuage concerns related to this issue, we conduct a variety of robustness checks to test the sensitivity of our results.

2.5 Results

2.5.1 Highest Grade Attained

First, we turn our attention to the models in Tables 2.1a and 2.1b, in which we show the relationship between second–grade test scores and the highest grade attained by young adults in Senegal and Madagascar, respectively. It should be noted that the highest grade attained might be different from the number of years of schooling. This is because repetition of grades is quite common in the two countries in our sample, as it is the case in most African countries that follow the French educational model. The first column of Tables 2.1a and 2.1b display the results from OLS regressions using a single covariate, the composite French and math score from the second–grade posttest. As pointed out previously, the test score variables have been created using IRT; thus, the test score mean and standard deviation are close to zero and one, respectively. In Senegal, a second–grade composite test score one standard deviation above the mean is associated with an increase in the highest grade attained by around 1.64 years. In Madagascar, the corresponding coefficient is 0.99 (Table 2.1b, column 1). Both of these coefficients are significant at the one percent level.

In columns 2 through 5, we introduce school fixed effects into the model. School fixed effects account for all time–invariant school characteristics and, hence, control for school–specific factors that impact young adult–life cognitive scores. As noted earlier, since each community had one school, the school fixed effects can also be thought of as community–level fixed effects. The coefficients in column (2) change little relative to column (1), remaining significant at the one per cent level in both countries.

In columns (3) and (4), we add a series of household and individual covariates. They do not lead to noteworthy changes in the coefficient of the second–grade test scores in either country and the test score coefficient remains strongly statistically significant (at 1 percent level) in these specifications. The fathers education level has a positive relationship with grade attainment in Madagascar, with the mothers education having a modest additional effect. In Senegal, the average level of parents education is low, so we use dummies for whether each parent has any education, instead of using a continuous measure[19]. The results for Senegal indicate that parents education has a positive, albeit statistically insignificant, relationship with grade attainment.

The second–grade household asset index, created using factor analysis, has a large positive and significant association with the highest grade attained in Senegal. We find that an increase of one unit in the asset index raises schooling by around 0.50 years in Senegal. In Madagascar, we do not see any significant effects of assets in early childhood. This might be because of the lower overall level of assets in households in Madagascar, as compared to those in Sene-

[19]In Senegal, mothers education is 1.3 years and fathers 2.7 years, on average. In Madagascar, mothers and fathers have 5.6 and 6.2 years of education, respectively.

gal[20]. It could also be that the coefficient estimate of the asset index is insignificant, because the regressions already control for parental education, which is an important determinant of living standards of the household[21]. We also tried adding interaction terms of the second–grade scores with assets and parents educationthese variables were not significant and therefore omitted from the specifications reported here.

In columns (4) and (5), we add the height of the cohort member into the model. As discussed in Case and Paxson, 2008a, Vogl, 2014, LaFave and Thomas, 2017, height is a proxy measure of childhood health and nutritional status, particularly, as affected by in utero and early childhood inputs. Results indicate that, in both countries, the coefficient on second–grade test scores is largely unaffected by the inclusion of height. In Senegal, height has a significant positive relationship with highest grade attained, whereas in Madagascar the coefficient albeit positive, is much smaller in magnitude and not significant. Our results indicate that being 1 cm taller is associated with an increase of 0.04 years of schooling in Senegal. The fact that we find that the relationship of early childhood and adult test scores is not affected by the inclusion of the height variable shows that early–life health and human capital (measured by test scores) have independent effects on adult outcomes[22]. This is largely consistent with the current literature, which found a statistically significant effect of height on human capital formation (Persico et al., 2004, Case and Paxson, 2008a, 2010, Spears, 2012).

[20]Summary statistics on the number of assets owned are available from the authors by request.

[21]The coefficient of the asset index is significant when we remove the parental education variable.

[22]In the case of Madagascar, we have also run the model controlling for ethnicity, which should be correlated with height, given that there is a mix of ethnic groups that are both Asian and African in origin. The results remain similar, when ethnicity is controlled for. Results are available from the authors by request.

As explained earlier, the second–grade test score suffers from idiosyncratic measurement error problems, which could lead to a biased estimate of its coefficient. In column (5) of Tables 2.1a and 2.1b, we report the results from the IV strategy that corrects for this measurement error by instrumenting the composite test score taken at the end of second grade with the score on the test administered at the beginning of second grade. The F–statistic for the excluded instrument (labeled widstat) is 254.8 in Senegal and 82.7 in Madagascar, which is well above the conventional threshold of 10 considered for weak instruments.

The magnitudes of the IV coefficients of the impact of composite test scores on grade attainment are similar in both countries, 1.38 and 1.29 for Senegal and Madagascar, respectively. Thus, the IV results portray a consistent narrative of a significant positive relationship between second–grade test scores and educational achievement later in life. It is not clear, however, why the idiosyncratic measurement error correction matters more in Madagascar than in Senegal. This may reflect the fact that the measurement error correction seems to work better with the pretest conducted in Madagascar than it does with that in Senegal, or that there might have been greater measurement error in Madagascar to start with. In addition, the difference in the correction from the IV can also stem from the fact that the questions in the pretests (used as the instrument) differed across the two countries.

2.5.2 Test Scores

In Tables 2.2a and 2.2b, we estimate the relationship between second–grade test scores and adult composite French and math test scores (columns 1 and 2), as well as math (columns 3 and 4) and French separately (columns 5 and 6). The findings in these tables are consistent with the results discussed previously in terms of grade attainment: second–grade cognitive ability has a strong and persistent association with later–life skills. More specifically, columns 1 and 2 show evidence of a robust positive and statistically significant relationship between second–grade skills and later–life composite French and math scores in Senegal and Madagascar. Consistent with the attainment models, the magnitude on the test score parameter rises in Madagascar when we adjust for measurement error using IV regressions (column 2). The results in the IV model in column 2 of Tables 2.a and 2.b show that a 1 standard deviation increase in composite scores in second grade is associated with higher adult composite scores by 0.27 and 0.32 standard deviation in Senegal and Madagascar, respectively. The results for Senegal are statistically significant at the 1 percent level, whereas the Madagascar results are significant at the 5 percent level.

As expected, results in Table 2.2a suggest that, in Senegal, the assets of the household when the cohort member was in the second grade is positively and significantly associated with later–life cognition, even after controlling for parents education. A standard deviation increase in the asset index is associated with an increase in the composite test score of 0.14 standard deviation. In Madagascar, although the asset index coefficient is positive in all the models, it is not statistically significant. Mothers education has a positive and marginally significant relationship with the composite test score in Madagascar.

We observe similar patterns in the results in columns 3 through 6 in Tables 2.2a and 2.2b, in which the individual scores on 2012 math and French tests are modeled separately. Childhood skills have a statistically significant positive association with both adult math and adult French scores in Senegal and Madagascar. However, the coefficient estimate is far stronger in the case of math than for French[23]. Overall, the results describe a situation in which childhood test scores are strongly and persistently associated with later–life human capital outcomes. These relationships hold even after the addition of control variables, the introduction of an IV strategy to correct for measurement error, and the use of school fixed effects. Thus, we provide persuasive evidence of the importance of better performance on tests in as early as second grade on adult human capital outcomes.

2.5.3 Heterogeneity tests

Differential results of French and math scores

In order to explore another dimension of the relationships discussed above, we replicate the regressions using a slightly modified empirical strategy. Instead of using the composite math and French scores in childhood, we enter the math and French scores separately as independent variables in different regression models. In the corresponding IV regressions, we use the French (math) test administered before second grade as an instrument for the French (math) test scores taken at the end of second grade[24].

We are motivated to do so because math and French tests potentially capture different types of abilities. Previous research has found that math skills in child-

[23]Using the z–scores of the percentage of correct answers as a dependent variable yields very similar results to the ones presented here. This is due to the fact that the z–scores and the IRT scores are very highly correlated. The results are available from the authors by request.

[24]The results are qualitatively similar if we were to use the pretest as the independent variable and the posttest as the instrument. The results are available from the authors by request.

hood are stronger predictors of later–life skills than language skills, although this evidence is from predominantly English–speaking countries (Duncan et al., 2007, 2011). Our results show that there is a strong and positive association of highest grade attained with second-grade math scores in both countries (Tables 2.3a and 2.3b, column 1). This is consistent with our main results, based on using the composite score (Tables 2.1a and 2.1b). In Senegal, a standard deviation increase in the second–grade math score is associated with an increase in highest grade attained by 1.4 years, whereas the corresponding coefficient estimate in Madagascar is 1.18 years (column 2 of Tables 2.3a and 2.3b, respectively).

In the first two columns of Tables 2.4a and 2.4b, we present similar evidence on the relationship between second–grade French scores and later–life grade attainment. The relationship between a 1 standard deviation increase in the second-grade French test score and highest grade attained is around 1.6 and 1.7 years in Senegal and Madagascar, respectively (column 2 in Tables 2.4a and 2.4b). These coefficient estimates are similar to the results that use the composite test score (Tables 2.1a and 2.1b).

In columns 3 to 8 of Tables 2.3a, 2.3b, 2.4a, and 2.4b, we run similar models, but this time the dependent variables are scores on the composite, math, and French tests in the second grade, respectively. In Senegal, the French score has a statistically significant association with all cognition outcomes, while the magnitudes are similar to the coefficients we get from the main specifications (Tables 2.1a and 2.1b). The results with second–grade French tests are relatively weaker in Madagascar (Table 2.4b). These results suggest that childhood math scores are stronger predictors of later–life math scores of later–life French scores, particularly in Senegal. Additionally, childhood math scores predict later–life

French scores better than childhood French scores do, in the case of Madagascar. For Senegal, adult French scores are equally well predicted by both childhood French and math scores. In sum, there is some evidence that the math scores are driving the strong relationship between composite scores (math and French) and later–life outcomes in Madagascar (Table 2.1b). We also note that the importance of other background characteristics is similar in the models when math, French, and composite scores from early childhood are used as covariates in the model.

Gender Differences

We also explore whether there are any gender differences in the relationship between childhood skills and later-life outcomes by running separate models for boys and girls. The results in Tables 2.5a and 2.5b indicate that the test score coefficient differs in magnitude between girls and boys in both countries. Across all outcomes in both countries, the coefficient for the childhood test score is consistently higher for girls than for boys. This gender difference is especially pronounced in Senegal. We conduct a t–test to check for the equality of the test score coefficients in the male and female regressions and reject the null of equality of coefficients in one case in Madagascar and in two cases in Senegal. The evidence that childhood performance is potentially more persistent in its impact on later–life cognitive ability for girls implies larger negative consequences for girls who fall behind in early grades. In other words, catching up from early cognitive deficits may be harder for girls, as compared to boys in both Madagascar and Senegal.

Height Differences

We next divide the sample into two groups based on whether the cohort members height falls above or below the median gender–specific height in each country (Tables 2.6a and 2.6b). In Senegal, the second–grade test score coefficient in the grade attainment models for the below–median group (relatively shorter) is greater than the coefficient in the above–median group (relatively taller). These differences, however, are not statistically significant. In Madagascar, the patterns are similar, but, again, we do not find statistically significant differences in the coefficient on test scores of the two groups of relatively shorter/taller cohort members. Taken together, these results provide some suggestive evidence that there is greater persistence in test scores from childhood to adulthood among shorter individuals. To the extent that shorter and less healthy cohort members are not only more vulnerable in childhood, but also have a higher persistence in their poor performance over time, the result suggests early deficits are unlikely to be overcome unless concerted investment/effort is expended to correct them[25].

2.6 Robustness Checks

2.6.1 Accounting for Attrition and Sample Selection

The attrition rates and sample selection described in Section 2 might raise a concern that our results could be driven by some form of sample selection. There-

[25]Due to the large ethnic diversity in Madagascar, we also ran the model in Madagascar that controlled for ethnicity. The results are qualitatively similar and not reported here.

fore, in Appendix C, we investigate the robustness of our findings to adjustments for attrition and sample selection for the follow–up. Recall from our discussion above that only a subset of the communities was randomly selected for follow–up, but this selection process does not necessarily ensure that they are representative of the original communities. First, in Table A2.2 we test whether there are systematic differences between the sample of observations in the panel and the full sample of students at baseline. In the balance test, we include a number of school–level covariates, available from the PASEC surveys, to check whether the school environments differ across the full baseline sample and the panel sample. We only find modest differences, mostly in the Senegalese sample. To control for these differences, we run school fixed effects models, which account for all time–invariant, school–specific characteristics.

In Tables A2.3 and A2.4, we compare other subsamples. Table A2.3 compares the means for observations in clusters chosen for follow-up with the full sample of students at baseline. We find that, for both countries, the results in Table A2.3 are similar to the ones that we found in Table A2.2, which are accounted for by the school fixed effects. However, we find no systematic mean differences within the clusters that were chosen for follow–up (Table CA2.4). That is, individuals reached and not reached within the communities that were chosen for the follow–up were very similar at baseline. To conclude, the balance checks presented in Tables A2.2–A2.4 show that, insofar as there are differences between the baseline PASEC sample and the panel subsample, they are driven by the geographical differences arising from the fact that only a subset of PASEC communities were included in the follow–up, and not due to differences between individuals that were reached and not reached within the selected communities. This finding applies for both countries.

Second, we estimate Inverse Probability Weighted (IPW) regressions, to account for these differences in the panel subsample and the full sample at baseline. These weights are obtained from a logistic regression and use a dummy variable denoting the probability of being in the panel sample as the dependent variable. This model contains a variety of household – and individual–level covariates from the full sample of the first round of data as covariates[26]. This specification is run using the baseline data, and the predicted probabilities are used as weights in the main regression to check their robustness to this adjustment. Table A2.5 replicate results of Tables 1.a and 1.b, columns 1–5, for Senegal and Madagascar, respectively, using the sample adjusted with the inverse probability weights. A comparison of Tables 2.1a and A2.5 shows that the results for the highest grade obtained in Senegal are consistent in sign, significance, and value, even after the aforementioned attrition adjustment. A comparison of 2.1b and A2.5 shows that the magnitudes are very similar in the case of Madagascar as well, and that most of the statistical significance levels also remain the same. The pattern of the results is similar for the test score outcome variables (comparing Tables 2.2a and 2.2b with A2.5, columns 6–8). The fact that the results do not change considerably when adjustments are made for attrition and sample selection demonstrates the robustness of this relationship across the two countries, despite small differences in the samples between the full original sample from the PASEC surveys in the mid–1990s, and the long panel of cohort members that we were able to track over the 17–year interval in Senegal.

We do a similar analysis to study attrition from within the clusters chosen for follow–up. More specifically, we run the IPW regressions (described above)

[26]Namely, the test scores of the second grade, gender, asset index, a school–level infrastructure index constructed with factor analysis, and the education level of the teacher. Observations with a missing weight have been given the average IPW weight.

based on the weights from a logistic model, which estimates the probability of being in the panel sample, conditional on being in a community that was chosen for the follow–up. The results are presented in Table A2.6 for Senegal and Madagascar, respectively. A comparison with Tables 2.1a and 2.1b shows that the results are robust to attrition in both countriesthe coefficient estimates of the grade attained are very similar in both magnitude and significance. Finally, columns 6–8 in Tables A2.6 confirm also that the coefficient estimates of the test score variables are similar to those in Tables 2.2a and 2.2b. They show a similar pattern of math scores being more persistent than French scores in both countries. Overall, we can conclude that even though the selection of only a subset of the communities for follow–up introduced some differences across the panel and the full PASEC sample at baseline, we find that our results are robust to this sample selection and also to accounting for the fact that the follow–up only included a subset of the children in the communities that were visited again.

2.6.2 Corrections based on Lewbel, 2012

In another robustness check, we complement our main IV strategy that is correcting for measurement error with a novel methodological approach. Lewbel, 2012 described an empirical framework in which the IV strategy exploits heteroskedasticity, in place of imposing the standard exclusion restrictions in the two–stage least squares framework. There are two main conditions that need to be satisfied to be able to apply this model: the presence of at least one exogenous variable in the structural equation and the heteroskedasticity of the error terms. This set of exogenous variables (Z) could be a subset of the independent variables (X) or could be the same as them. Under this method, one regresses each endogenous regressor on the set of exogenous variables. The residuals from these regressions are used along with the demeaned set of exogenous variables to construct generated instruments. This estimation framework is similar in nature to other approaches in which heteroskedasticity has been used as a source of identification (King et al., 1994, Heckman and Vytlacil, 1998, Sentana and Fiorentini, 2001 among others).

In our case, we present results using the pretest and the generated instruments as instrumental variables in our model[27]. The inclusion of an extra instrument allows us to conduct the Sargan–Hansen overidentification test. Under the null that the overidentifying restrictions are valid, the test has a chi–square distribution. We are unable to compute this statistic in our main tables, because the IV models are exactly identified, that is, the number of instruments is equal to the number of endogenous regressors. The usage of the Lewbel, 2012 method allows us to conduct this test as the generated instruments make the model overidentified. These results can be compared to the IV regression results in Tables 2.1a–2.2b, which use only the pretest as an instrument.

The results in Table 8 indicate that the addition of the generated instrument using this method does not significantly alter the IV results. This is especially the case in Senegal, where the coefficient estimate of the test score on all the outcomes remains relatively stable (as compared to Tables 2.1a and 2.2a) and statistically significant at the 1 percent level. The results for Madagascar lose a little bit of statistical significance but still retain the correct sign and are of a similar magnitude as the IV results in Tables 2.1b and 2.2b. In addition, the J–statistic p–value shows that the null hypothesis of valid overidentifying restrictions is valid for all outcomes across both countries. Therefore, we conclude that our

[27]We do not present the models with only the generated instruments.

main IV specification, employed to correct for measurement error, is robust to the Lewbel, 2012 instrumental variables strategy.

2.7 Conclusions

We find persuasive evidence of a strong association of childhood academic skills with those measured in early adult life in two francophone sub–Saharan African countries, Senegal and Madagascar. Using a production function framework for human capital, we find that composite math and French test scores, measured in the second grade, have large and significant positive associations with the highest grade attained, as well as math and French test scores in young adulthood in both countries. This enduring relationship is stronger in the case of Senegal, as compared to Madagascar; we also find that childhood math scores are stronger predictors of later–life cognitive outcomes, as compared to childhood French scores. This finding is consistent with results reported elsewhere that indicate certain types of abilities in childhood are more important in predicting human capital outcomes later in life.

We also explore whether lifetime cumulative health, measured using adult height, is significantly associated with adult human capital, and whether its inclusion in the models affects the strength of the relationship between second–grade test scores and adult cognition. We only find statistically significant coefficient estimates of height in Senegal, although in Madagascar, the sign and magnitude of the coefficient is quite similar. Despite the inclusion of height, the aforementioned relationships between childhood test scores and later– life schooling and skills are persistent and found to operate through independent channels.

The results we report are robust to the addition of other childhood inputs, namely parental education and asset levels when the cohort member was in the second grade, as well as school fixed effects. Parental inputs have an independent relationship with early adulthood outcomes in both countries. Household assets measured in second grade have a significant positive association with adulthood outcomes, even controlling for early test scores and other variables in Senegal and whereas in Madagascar, parents education matters more.

We also run a series of heterogeneity tests and find that there are larger negative consequences for girls who fall behind in early grades. We similarly find that shorter and less healthy cohort members have a higher persistence in their poor cognitive performance over time. In other words, catching up from early cognitive deficits may be harder for girls and unhealthy children. In contrast, low levels of assets early in life do not seem to imply that children from relatively richer households are better able to sustain their better performance in second grade later into life.

Additionally, we discuss challenges that arose, due to the ambitions of examining test scores over a span of more than 15 years, including issues of attrition, as well the potential for measurement error. By employing techniques, such as estimating inverse probability weighted regressions, and employing the Lewbel, 2012 instrumental variable method, we show that our results are robust to these potential issues.

While we do not directly address policies to improve cognitive outcomes of young adults, our results imply that childhood academic skills are a powerful predictor of young adulthood human capital outcomes. In turn, this implies that policies should target preschool–aged children who are lagging behind other children in terms of their skills and health status, and that such interventions are particularly important for young girls (and shorter individuals) who seem less able to catch up from early academic disadvantage.

2.8 Tables

Table 2.1a: Highest grade completed as a function of childhood composite French and math scores – Senegal

	(1) No School FE OLS	(2) School FE OLS	(3) School FE OLS	(4) School FE OLS	(5) School FE IV
Second Grade Composite Score	1.645***	1.783***	1.728***	1.695***	1.380***
	(0.185)	(0.207)	(0.194)	(0.195)	(0.310)
Height (in 2012 cms)				0.042*	0.045**
				(0.023)	(0.021)
Assets in second grade			0.495*	0.487*	0.549**
			(0.285)	(0.285)	(0.271)
Mother's Education (dummy)			0.616	0.473	0.459
			(0.582)	(0.596)	(0.559)
Father's Education (dummy)			0.331	0.285	0.285
			(0.516)	(0.516)	(0.481)
Age in 2012 (years)			-0.496***	-0.498***	-0.497***
			(0.08)	(0.08)	(0.074)
Female			-0.151	0.262	0.227
			(0.330)	(0.406)	(0.380)
Total Obs.	447	447	447	447	447
R-squared	0.143	0.349	0.413	0.419	0.235
F-stat (instrument)					254.8

Second grade denotes the year 1995–96 in the case of Senegal and 1997–98 in Madagascar. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The F–stat denotes the Kleibergen-Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.1b: Highest grade completed as a function of childhood composite French and math scores – Madagascar

	(1) No School FE OLS	(2) School FE OLS	(3) School FE OLS	(4) School FE OLS	(5) School FE IV
Second Grade Composite Score	0.993***	0.716***	0.666***	0.665***	1.285***
	(0.193)	(0.273)	(0.232)	(0.232)	(0.462)
Height (in 2012 cms)				0.019	0.019
				(0.02)	(0.018)
Assets in second grade			-0.063	-0.059	-0.119
			(0.246)	(0.248)	(0.242)
Mother's Education			0.09*	0.087	0.077
			(0.053)	(0.053)	(0.05)
Father's Education			0.145***	0.141***	0.137***
			(0.049)	(0.049)	(0.045)
Age in 2012 (years)			-0.707***	-0.711***	-0.733***
			(0.123)	(0.123)	(0.113)
Female			-0.267	-0.116	-0.096
			(0.301)	(0.327)	(0.300)
Total Obs.	333	333	333	333	333
R–squared	0.085	0.366	0.496	0.498	0.209
F–stat (instrument)					82.7

Second grade denotes the year 1995–96 in the case of Senegal and 1997–98 in Madagascar. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The F–stat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.2a: Adult test scores as a function of childhood composite French and math scores – Senegal

	(1) Composite OLS	(2) Composite IV	(3) Math OLS	(4) Math IV	(5) French OLS	(6) French IV
Second Grade Composite Score	0.362***	0.269***	0.625***	0.556***	0.307***	0.210***
	(0.051)	(0.070)	(0.08)	(0.117)	(0.048)	(0.068)
Height (in 2012 cms)	0.007	0.008	0.012	0.012	0.006	0.007
	(0.006)	(0.005)	(0.008)	(0.008)	(0.006)	(0.005)
Assets in second grade	0.123*	0.141**	0.20*	0.213**	0.134**	0.152**
	(0.067)	(0.063)	(0.105)	(0.101)	(0.065)	(0.061)
Mother's Education (dummy)	-0.078	-0.072	-0.068	-0.071	-0.024	-0.018
	(0.145)	(0.135)	(0.212)	(0.199)	(0.138)	(0.128)
Father's Education (dummy)	0.016	0.016	0.028	0.028	0.045	0.046
	(0.128)	(0.118)	(0.177)	(0.165)	(0.128)	(0.118)
Age in 2012 (years)	-0.061***	-0.061***	-0.105***	-0.105***	-0.057***	-0.058***
	(0.021)	(0.019)	(0.03)	(0.028)	(0.02)	(0.018)
Female	0.072	0.06	-0.069	-0.07	0.101	0.089
	(0.106)	(0.098)	(0.153)	(0.142)	(0.103)	(0.095)
Total Obs.	381	381	447	447	381	381
R-squared	0.351	0.166	0.327	0.193	0.342	0.140
F-stat (instrument)		232.9		254.8		232.9

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.2b: Adult test scores as a function of childhood composite French and math scores – Madagascar

	(1) Composite OLS	(2) Composite IV	(3) Math OLS	(4) Math IV	(5) French OLS	(6) French IV
Second Grade Composite Score	0.146**	0.316**	0.154**	0.349**	0.127*	0.26*
	(0.064)	(0.134)	(0.07)	(0.139)	(0.068)	(0.142)
Height (in 2012 cms)	0.005	0.005	-0.001	-0.001	0.005	0.005
	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Assets in second grade	0.064	0.052	0.088	0.074	0.019	0.009
	(0.07)	(0.065)	(0.067)	(0.063)	(0.082)	(0.077)
Mother's Education	0.026*	0.023*	0.024	0.021	0.03**	0.028**
	(0.014)	(0.013)	(0.016)	(0.015)	(0.013)	(0.012)
Father's Education	0.017	0.016	-0.002	-0.004	0.036***	0.035***
	(0.012)	(0.012)	(0.014)	(0.013)	(0.012)	(0.011)
Age in 2012 (years)	-0.061***	-0.061***	-0.105***	-0.105***	-0.057***	-0.058***
	(0.021)	(0.019)	(0.03)	(0.028)	(0.02)	(0.018)
Female	-0.029	-0.025	-0.150	-0.143	0.07	0.073
	(0.101)	(0.092)	(0.102)	(0.095)	(0.107)	(0.097)
Total Obs.	310	310	318	318	312	312
R–squared	0.490	0.118	0.377	0.071	0.529	0.133
F–stat (instrument)		57.8		60.01		57.39

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects. All test scores are constructed using country-specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.3a: Adult test scores as a function of childhood Math scores – Senegal

	(1) Grade OLS	(2) Grade IV	(3) Composite OLS	(4) Composite IV	(5) Math OLS	(6) Math IV	(7) French OLS	(8) French IV
Second Grade Math Score	1.383***	1.406***	0.284***	0.267***	0.532***	0.566***	0.223***	0.211***
	(0.202)	(0.330)	(0.052)	(0.077)	(0.079)	(0.125)	(0.048)	(0.073)
Height (in 2012 cms)	0.041*	0.041*	0.007	0.007	0.011	0.011	0.007	0.007
	(0.023)	(0.022)	(0.006)	(0.005)	(0.008)	(0.008)	(0.006)	(0.005)
Assets in second grade	0.575**	0.571**	0.135**	0.138**	0.228**	0.222**	0.147**	0.150**
	(0.285)	(0.268)	(0.068)	(0.064)	(0.104)	(0.100)	(0.067)	(0.062)
Mother's Education (Dummy)	0.560	0.563	-0.055	-0.055	-0.033	-0.029	-0.005	-0.005
	(0.616)	(0.576)	(0.146)	(0.134)	(0.212)	(0.197)	(0.140)	(0.129)
Father's Education (Dummy)	0.309	0.309	0.011	0.012	0.037	0.038	0.042	0.042
	(0.531)	(0.493)	(0.131)	(0.120)	(0.183)	(0.170)	(0.131)	(0.120)
Age in 2012 (years)	-0.500***	-0.501***	-0.062***	-0.062***	-0.106***	-0.106***	-0.058***	-0.058***
	(0.082)	(0.076)	(0.021)	(0.019)	(0.03)	(0.028)	(0.02)	(0.018)
Female	0.397	0.403	0.098	0.093	-0.014	-0.006	0.118	0.115
	(0.414)	(0.392)	(0.107)	(0.101)	(0.153)	(0.145)	(0.104)	(0.097)
Total Obs.	447	447	381	381	447	447	381	381
R–squared	0.396	0.210	0.323	0.138	0.310	0.173	0.312	0.111
F–stat (instrument)		191.9		169.3		191.9		169.3

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.3b: Adult test scores as a function of childhood composite Math scores – Madagascar

	(1) Grade OLS	(2) Grade IV	(3) Composite OLS	(4) Composite IV	(5) Math OLS	(6) Math IV	(7) French OLS	(8) French IV
Second Grade Math Score	0.617***	1.179**	0.161***	0.356**	0.190***	0.336*	0.125***	0.355**
	(0.194)	(0.471)	(0.055)	(0.173)	(0.061)	(0.180)	(0.057)	(0.173)
Height (in 2012 cms)	0.018	0.017	0.005	0.005	-0.002	-0.002	0.005	0.005
	(0.02)	(0.019)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Assets in second grade	-0.045	-0.09	0.065	0.053	0.088	0.079	0.021	0.007
	(0.245)	(0.238)	(0.069)	(0.064)	(0.066)	(0.060)	(0.082)	(0.079)
Mother's Education	0.092*	0.087*	0.027*	0.025*	0.024	0.023	0.031*	0.029*
	(0.052)	(0.049)	(0.014)	(0.013)	(0.016)	(0.015)	(0.013)	(0.013)
Father's Education	0.140***	0.135***	0.017	0.015	-0.003	-0.004	0.036***	0.033***
	(0.049)	(0.045)	(0.012)	(0.012)	(0.014)	(0.013)	(0.012)	(0.012)
Age in 2012 (years)	-0.715***	-0.739***	-0.134***	-0.147***	-0.123***	-0.133***	-0.101***	-0.117***
	(0.122)	(0.114)	(0.036)	(0.035)	(0.037)	(0.037)	(0.039)	(0.039)
Female	-0.085	-0.038	-0.02	-0.005	-0.14	-0.128	0.076	0.094
	(0.329)	(0.306)	(0.101)	(0.094)	(0.103)	(0.095)	(0.107)	(0.099)
Total Obs.	333	333	310	310	318	318	312	312
R–squared	0.500	0.210	0.496	0.112	0.388	0.093	0.531	0.099
F–stat (instrument)		51.28		45.16		45.84		44.94

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.4a: Adult test scores as a function of childhood composite French scores – Senegal

	(1) Grade OLS	(2) Grade IV	(3) Composite OLS	(4) Composite IV	(5) Math OLS	(6) Math IV	(7) French OLS	(8) French IV
Second Grade French Score	1.652***	1.592***	0.363***	0.333***	0.595***	0.650***	0.327***	0.251***
	(0.203)	(0.409)	(0.052)	(0.099)	(0.081)	(0.157)	(0.050)	(0.096)
Height (in 2012 cms)	0.048**	0.049**	0.008	0.008	0.014*	0.014*	0.007	0.007
	(0.023)	(0.021)	(0.006)	(0.005)	(0.008)	(0.008)	(0.006)	(0.005)
Assets in second grade	0.511*	0.522*	0.136**	0.140**	0.211**	0.201*	0.141**	0.153**
	(0.293)	(0.286)	(0.069)	(0.064)	(0.108)	(0.106)	(0.066)	(0.062)
Mother's Education (Dummy)	0.439	0.437	-0.083	-0.081	-0.081	-0.080	-0.030	-0.024
	(0.600)	(0.558)	(0.150)	(0.137)	(0.220)	(0.204)	(0.140)	(0.129)
Father's Education (Dummy)	0.220	0.222	0.013	0.013	0.004	0.002	0.042	0.044
	(0.516)	(0.479)	(0.128)	(0.117)	(0.178)	(0.164)	(0.127)	(0.117)
Age in 2012 (years)	-0.499***	-0.499***	-0.062***	-0.062***	-0.105***	-0.106***	-0.059***	-0.059***
	(0.081)	(0.075)	(0.021)	(0.019)	(0.031)	(0.029)	(0.02)	(0.018)
Female	0.033	0.035	0.031	0.030	-0.153	-0.154	0.066	0.065
	(0.414)	(0.386)	(0.108)	(0.099)	(0.156)	(0.145)	(0.104)	(0.095)
Total Obs.	447	447	381	381	447	447	381	381
R–squared	0.405	0.221	0.343	0.163	0.308	0.171	0.344	0.148
F–stat (instrument)		121.4		101.3		121.4		101.3

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.4b: Adult test scores as a function of childhood French scores – Madagascar

	(1) Grade OLS	(2) Grade IV	(3) Composite OLS	(4) Composite IV	(5) Math OLS	(6) Math IV	(7) French OLS	(8) French IV
Second Grade French Score	0.424	1.696	0.047	0.303	0.006	0.420*	0.083	0.208
	(0.265)	(0.876)	(0.075)	(0.215)	(0.081)	(0.232)	(0.077)	(0.222)
Height (in 2012 cms)	0.019	0.020	0.005	0.005	-0.002	-0.001	0.005	0.005
	(0.02)	(0.019)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Assets in second grade	-0.027	-0.122	0.072	0.059	0.099	0.077	0.024	0.017
	(0.248)	(0.254)	(0.071)	(0.067)	(0.069)	(0.067)	(0.083)	(0.077)
Mother's Education	0.092*	0.075	0.028*	0.024*	0.026*	0.021	0.031**	0.030**
	(0.053)	(0.053)	(0.014)	(0.013)	(0.016)	(0.016)	(0.014)	(0.013)
Father's Education	0.144***	0.141***	0.018	0.017	-0.001	-0.003	0.036***	0.033***
	(0.049)	(0.047)	(0.013)	(0.012)	(0.014)	(0.013)	(0.012)	(0.011)
Age in 2012 (years)	-0.692***	-0.704***	-0.123***	-0.127***	-0.109***	-0.116***	-0.094***	-0.095***
	(0.126)	(0.117)	(0.036)	(0.034)	(0.037)	(0.035)	(0.039)	(0.036)
Female	-0.153	-0.201	-0.034	-0.046	-0.156	-0.168*	0.063	0.058
	(0.333)	(0.312)	(0.102)	(0.092)	(0.103)	(0.097)	(0.107)	(0.097)
Total Obs.	333	333	310	310	318	318	312	312
R–squared	0.488	0.156	0.481	0.088	0.366	0.006	0.525	0.130
F–stat (instrument)		31.73		25.56		26.65		25.19

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.5a: Gender Heterogeneity – Senegal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Edu Years	Edu Years	Composite	Composite	Math	Math	French	French
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Second Grade Composite Score	2.264***	1.231***	0.507***	0.259***	0.470***	0.201***	0.866***	0.578***
	(0.538)	(0.401)	(0.121)	(0.088)	(0.106)	(0.090)	(0.196)	(0.140)
Total Obs.	188	259	161	220	161	220	188	259
R–squared	0.344	0.184	0.273	0.156	0.279	0.125	0.256	0.183
F–stat (instrument)	67.03	180.8	66.58	149.3	66.58	149.3	67.03	180.8
P–value of diff		0.123		0.097		0.054		0.232

Table 2.5b: Gender Heterogeneity – Madagascar

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Edu Years	Edu Years	Composite	Composite	Math	Math	French	French
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Second Grade Composite Score	2.448***	0.586	0.620**	0.221	0.559*	0.161	0.685**	0.282
	(0.660)	(0.591)	(0.283)	(0.178)	(0.300)	(0.172)	(0.271)	(0.200)
Total Obs.	179	154	164	146	165	147	170	148
R–squared	0.216	0.201	0.114	0.142	0.153	0.144	0.035	0.079
F–stat (instrument)	41.94	30.97	20.45	25.94	20.58	25.17	24	25.27
P–value of diff.		0.036		0.233		0.251		0.232

All coefficients are from the second stage of IV regressions. Second grade denotes 1995—96 in the case of Senegal and 1997—98 in Madagascar. All the specifications include school–level fixed effects.Specifications same as in Tables 1,2 and 3, but not reported in this table (excluding female). Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.6a: Height Heterogeneity – Senegal

	(1) Edu Years	(2) Edu Years	(3) Composite	(4) Composite	(5) Math	(6) Math	(7) French	(8) French
	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med
Second Grade Composite Score	1.495***	1.638***	0.304***	0.319***	0.210*	0.259***	0.545***	0.648***
	(0.437)	(0.385)	(0.116)	(0.105)	(0.118)	(0.098)	(0.186)	(0.149)
Total Obs.	234	213	198	183	198	183	234	213
R–squared	0.313	0.275	0.183	0.236	0.156	0.232	0.194	0.219
F–stat (instrument)	100.3	156.1	84.57	156.3	84.57	156.3	100.3	156.1
P–value of diff		0.807		0.924		0.751		0.667

Table 2.6b: Height Heterogeneity – Madagascar

	(1) Edu Years	(2) Edu Years	(3) Composite	(4) Composite	(5) Math	(6) Math	(7) French	(8) French
	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med
Second Grade Composite Score	0.983**	1.105	0.22	0.443**	0.125	0.452*	0.312	0.476***
	(0.471)	(0.829)	(0.168)	(0.201)	(0.151)	(0.254)	(0.209)	(0.181)
Total Obs.	172	161	160	150	161	151	164	154
R–squared	0.192	0.225	0.08	0.136	0.12	0.117	0.029	0.09
F–stat (instrument)	55.84	25.34	30.63	23.05	30.68	22.93	33.28	23.15
P–value of diff.		0.898		0.393		0.268		0.554

All coefficients are from the second stage of IV regressions. Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects.Specifications same as in Tables 1,2 and 3, but not reported in this table (excluding female). Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.7a: Assett Heterogeneity – Senegal

	(1) Edu Years	(2) Edu Years	(3) Composite	(4) Composite	(5) Math	(6) Math	(7) French	(8) French
	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med
Second Grade Composite Score	1.820***	0.893*	0.300*	0.180	0.221*	0.143	0.707***	0.476***
	(0.513)	(0.467)	(0.125)	(0.113)	(0.121)	(0.112)	(0.202)	(0.168)
Total Obs.	224	223	190	191	190	191	224	223
R–squared	0.294	0.145	0.212	0.089	0.184	0.061	0.254	0.136
F–stat (instrument)	81.96	128.4	74.9	117.7	74.9	128.4	81.96	128.4
P–value of diff		0.181		0.477		0.636		0.379

Table 2.7b: Asset Heterogeneity – Madagascar

	(1) Edu Years	(2) Edu Years	(3) Composite	(4) Composite	(5) Math	(6) Math	(7) French	(8) French
	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med	Above Med	Below Med
Second Grade Composite Score	0.983**	1.105	0.22	0.443**	0.125	0.452*	0.312	0.476***
	(0.471)	(0.829)	(0.168)	(0.201)	(0.151)	(0.254)	(0.209)	(0.181)
Total Obs.	172	161	160	150	161	151	164	154
R–squared	0.192	0.225	0.08	0.136	0.12	0.117	0.029	0.09
F–stat (instrument)	55.84	25.34	30.63	23.05	30.68	22.93	33.28	23.15
P–value of diff.		0.898		0.393		0.268		0.554

All coefficients are from the second stage of IV regressions. Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects.Specifications same as in Tables 1,2 and 3, but not reported in this table (excluding female). Heteroscedasticity–robust standard errors in parentheses. Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 2.8: Lewbel Corrections – Senegal & Madagascar

	Senegal				**Madagascar**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Edu Years	Composite	Math	French	Edu Years	Composite	Math	French
Second Grade Composite Score	1.338***	0.265***	0.570***	0.203***	1.023*	0.24*	0.239*	0.196
	(0.327)	(0.076)	(0.127)	(0.074)	(0.478)	(0.128)	(0.133)	(0.137)
Total Obs.	447	381	447	381	333	310	318	312
R–squared	0.234	0.165	0.196	0.139	0.221	0.133	0.086	0.142
F–stat (instrument)	41.75	38.07	42.32	38.07	13.55	6.25	3.82	6.49
J–statistic p–value	0.172	0.716	0.928	0.525	0.395	0.699	0.631	0.663

All these models are IV models where the instruments are the pretest score in second grade and the generated instrument based on the method described in Lewbel, 2012. These specifications contain school fixed effects. The mothers and fathers education variables in Senegal are dummy variables for whether they have any education or not. In Madagascar, those variables are based on the number of years of education they have. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table 2.9: Robustness Check – Average of Pre & Post-test – Senegal & Madagascar

	Senegal				Madagascar			
	(1) Edu Years	(2) Composite	(3) Math	(4) French	(5) Edu Years	(6) Composite	(7) Math	(8) French
Second Grade Average Composite Score	1.612***	0.328***	0.608***	0.271***	0.885***	0.206***	0.221***	0.175**
	(0.237)	(0.058)	(0.093)	(0.055)	(0.283)	(0.079)	(0.082)	(0.084)
Total Obs.	447	381	447	381	333	310	318	312
R–squared	0.390	0.321	0.302	0.316	0.502	0.494	0.382	0.531

The main independent variable of interest is the second-grade average of pretest and posttests. Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All the specifications include school–level fixed effects. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

2.9 Figures

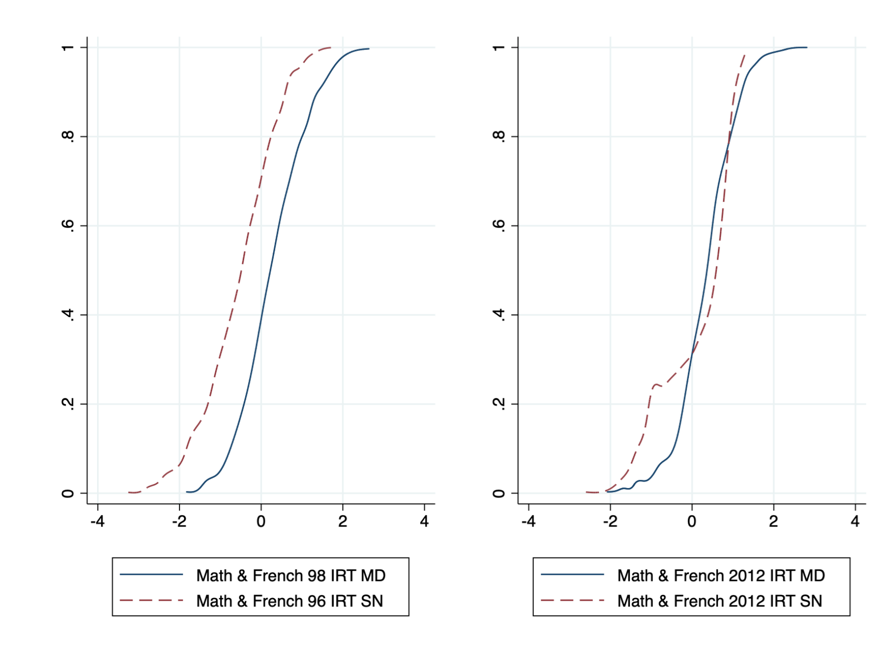

Figure 2.1: Cumulative distribution functions of composite scores

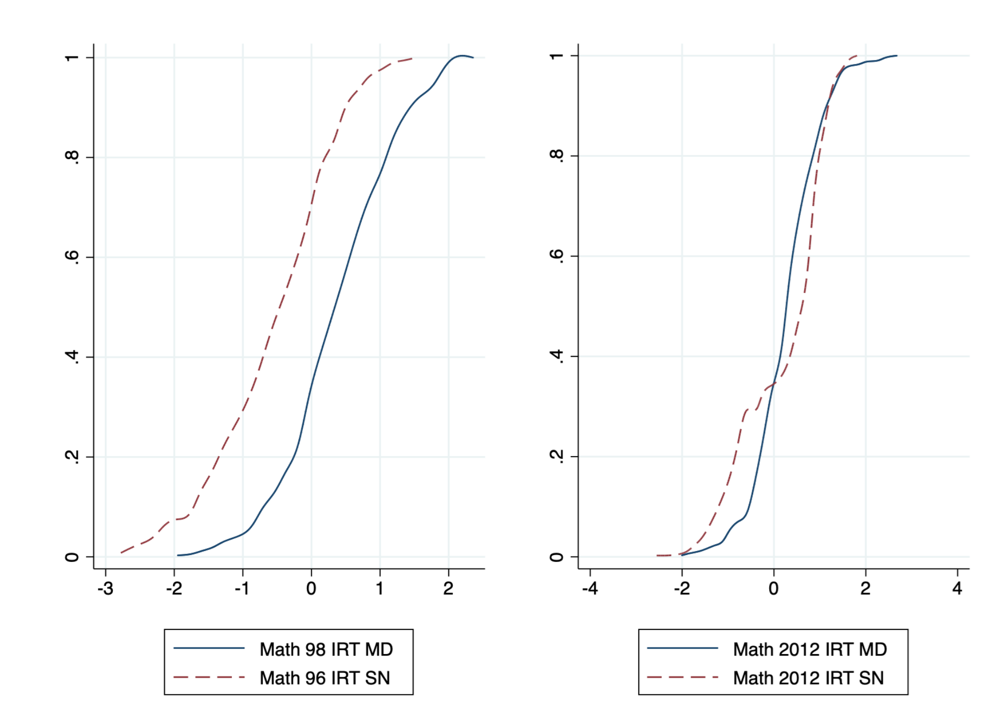

Figure 2.2: Cumulative distribution functions of math

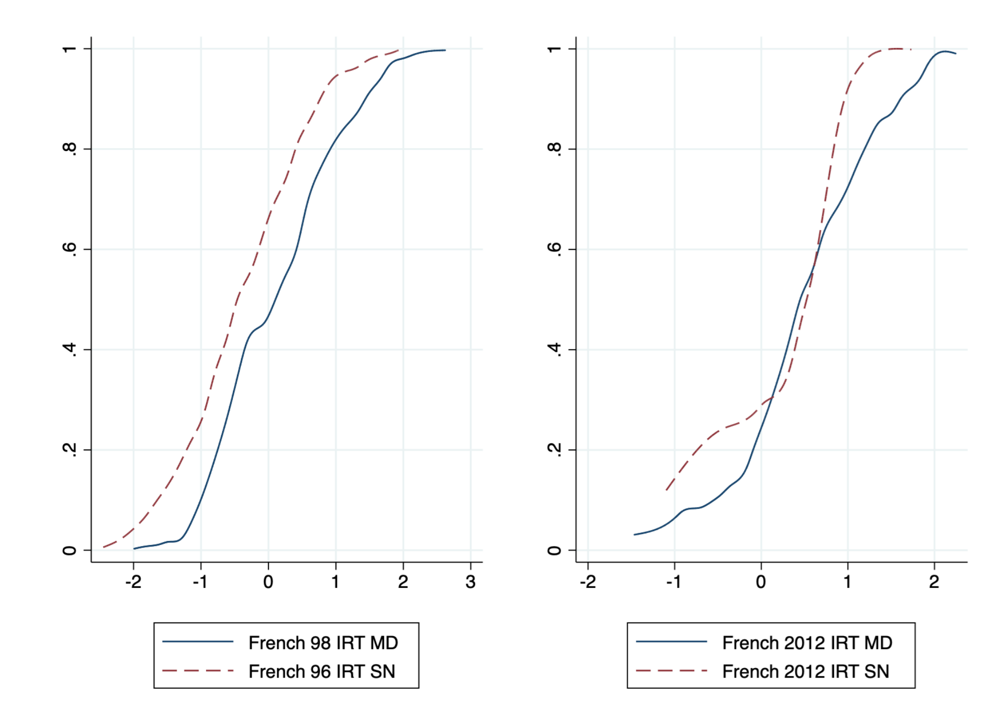

Figure 2.3: Cumulative distribution functions of French

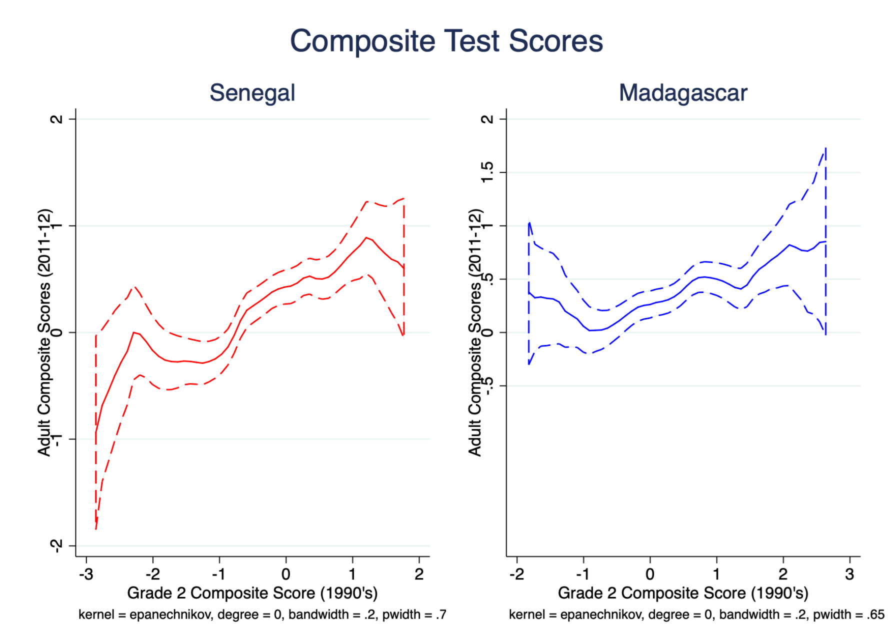

Figure 2.4: Learning progress curves composite scores

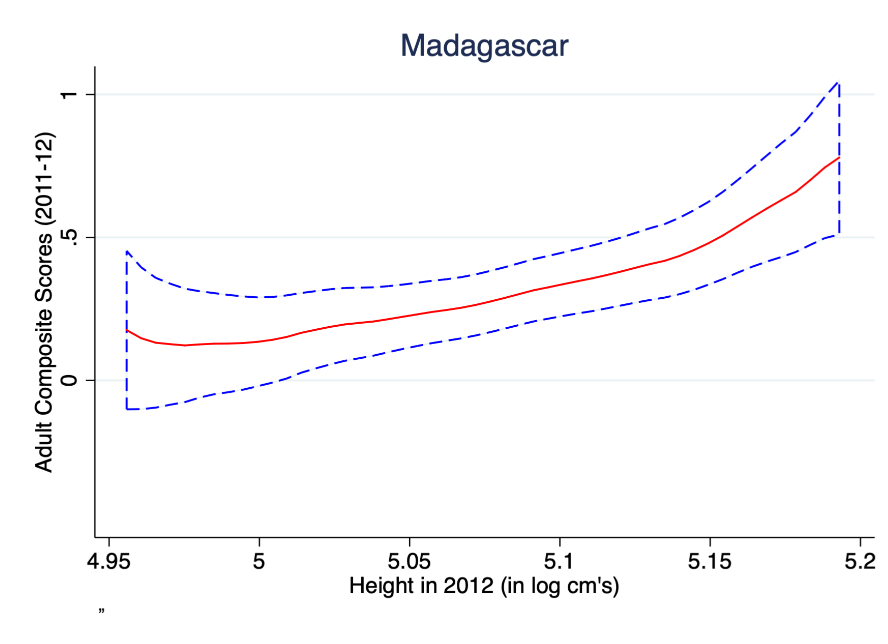

Figure 2.5: Height and composite test scores in 2012 – Madagascar

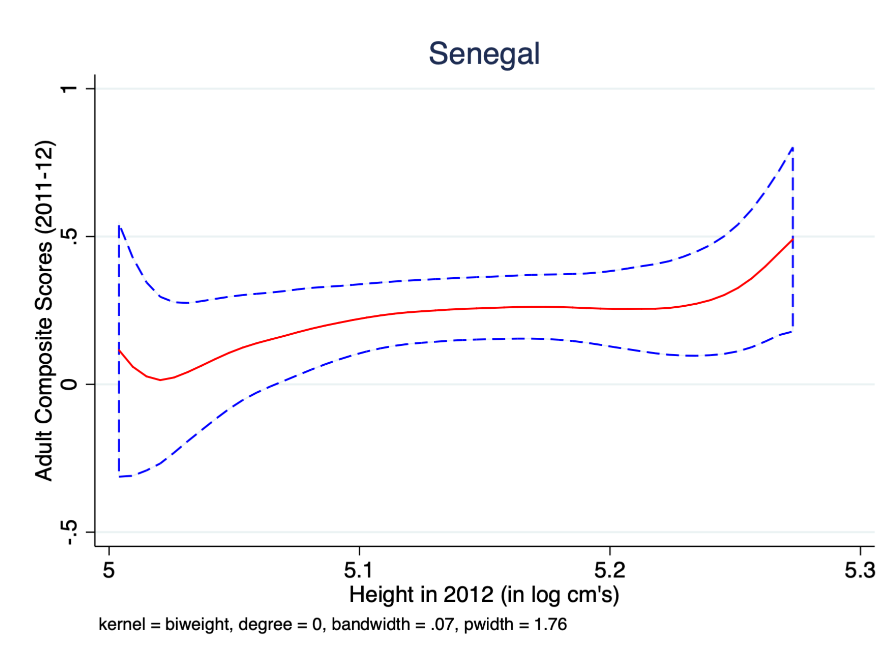

Figure 2.6: Height and composite test scores in 2012 – Senegal

CHAPTER 3

MARRIAGE AGE, SOCIAL STATUS AND INTERGENERATIONAL EFFECTS IN UGANDA

3.1 Introduction

Early or child marriage among girls, is a common practice in large parts of the developing world, especially in South Asia and Sub–Saharan Africa (SSA). It contributes greatly to the entrenchment of female disadvantages in these societies. More than 37% of marriages in SSA involve a child and in Uganda, which is the setting for the current study, 49% and 15% of women aged 20–49 years are married before the age of 18 and 15 years respectively (UBOS and ICF International, 2012[1]).

Studies in different contexts show that earlier marriage among women leads to a fall in female literacy and educational attainment (Field and Ambrus, 2008a, Sekhri and Debnath, 2014, Hicks and Hicks, 2015). The lower levels of education, associated with early marriage, might have further effects on employment and wages (as documented in Joshi and Schultz, 2007). In fact, Dahl, 2010 finds that women who marry earlier in life are 31 percentage points more likely to live in poverty when they are older. Early marriage can also have an effect on later life outcomes for the woman and her post–marital household through her preferences, and bargaining power (Glewwe, 1999, Christiaensen and Alderman, 2004, Banerji et al., 2017). These might in part be due to marriage market sorting on the kind of spouses (and households) that younger brides get married to, which might be worse than the average (Becker, 1973, Anderson, 2007).

[1] Accessed on 6 February 2018 – Link

Early marriage also poses large health risks to women and their children. Women who marry younger are less likely to be able to carry out fertility control, and thus avoid unwanted and terminated pregnancies (Raj et al., 2009). Early female marriage is also likely to be associated with early childbearing, which leads to a higher risk of maternal mortality and other pregnancy–related complications such as maternal anemia and pre–term labour (Clark et al., 2006, Nour, 2006). These complications have negative consequences for the health of both the mother and the child (Steer, 2000, Stoltzfus et al., 2004, Joanne, 2011, Goli et al., 2015). The intergenerational impacts might be mediated by women's bargaining power and their preferences regarding investment in children's human capital (Beegle et al., 2001, Majlesi, 2016).

In this paper, I explore the negative effects of early marriage among Ugandan women. In the main analysis, marriage age is conceptualized as a continuous variable where lower values imply marriage at a younger age. In a robustness check I check if the results are sensitive to alternative measures of early marriage. Various factors such as customs, beliefs and household characteristics could shape the age at which a woman gets married and also influence other circumstances of her life, and so marriage age is likely to be endogenous. I use an Instrumental Variable (IV) strategy to estimate the impacts of the timing of marriage for women – I use age at menarche to instrument for age of marriage (as used first by Field and Ambrus, 2008a). In many developing country contexts such as in Uganda, girls are married only after they reach puberty and so the onset of menarche is a binding constraint for the marriage of girls in these regions. I would thus expect marriage age to be strongly correlated with menarche age, with the relationship being positively signed. I show that such an association exists between the instrument and marriage age for the sample

I examine in my analysis. Furthermore, I argue that since menarche is largely biologically determined (Jahanfar et al., 2013, Adair, 2001), it is plausibly exogenous and affects later life outcomes only through its impact on marriage age. Having said that, in my empirical strategy I account for other factors that may affect menarche age (example – early life socioeconomic and nutrition inputs, shocks in infancy and altitude) or be affected by it (like age at first sex). I show that the effects I find remain robust to controlling for these additional factors.

I focus on four main categories of outcomes: schooling, work, health behaviors and child health, with the last category capturing the intergenerational effects of marriage age. My results suggest that early marriage leads to lower female educational attainment, literacy and labour force participation. Women who marry earlier also demonstrate poorer health behaviors related to the use of contraceptives, ante natal care and age at first birth. I find that children born to women who marry young have lower BMI and hemoglobin levels, and are at a higher risk of being anemic compared to children of women who marry at older ages. In exploring the mechanisms, apart from enhanced educational levels, through which the observed effects could have been mediated, I find evidence for the possible role of two other factors – greater female decision making power and positive assortative matching in the marriage market (that is, women who marry later tend to attract better quality spouses).

I demonstrate that the results I find are robust to various checks. First, while I measure marriage age as a continuous variable for my main analysis, I show that the results hold when I define early marriage to apply to those marrying below different age thresholds (18, 16 or 14 years). In another robustness check, I use the method proposed by Conley et al., 2012 to demonstrate that the coefficient estimates that I identify persist when I relax what is known as the exclusion restriction of an IV strategy, under which age at menarche is expected to shape the outcome(s) of interest only through marriage age and not through any other confounders. Analogous to this, I convey the stability of the results to increases in potential biases due to unobservable factors (as per Oster, 2014). I also demonstrate that the results remain unchanged when I use an alternative empirical model – probit instead of the Linear Probability Model (LPM) that I use for the main analysis.

This paper adds to the literature by estimating the impacts of early marriage among women in Uganda, a context in which this topic has not been explored before with rigorous empirical methods. Since child marriage is still widely practiced in Uganda, this is an important setting to study the topic in. The current analysis also adds to the literature that examines the long term consequences of early female marriage, for example on outcomes such as post–marriage labour force participation and decision making power. In addition, this is one of the few papers to explore the intergenerational health impacts of early marriage. Finally, this paper makes a valuable empirical contribution by demonstrating the external validity of an econometric methodology which has previously been used mostly in South Asian countries (Field and Ambrus, 2008a, Sekhri and Debnath, 2014).

3.2 Study Context: Uganda

Uganda is a low income country with a per capita Gross Domestic Product (GDP) of around USD 1300 (2010 data[2]) and an economy that is primarily agrarian. Uganda fares poorly in terms of health – the life expectancy at birth is 57 years, and the infant mortality rate is high at 61 per 1000 live births. Uganda's 2015 HDI value (a measures that reflects life expectancy, educational status and income) ranked it as 163rd in a list of 188 countries, making it one of the poorest performers across the globe. The HDI is often calculated separately for the female and male populations of a country and the ratio of these two figures (female HDI/male HDI) used to get a sense of the prevalence of gender disparity. Given that a GDI value below one implies gender inequality (skewed against women), Uganda's GDI of 0.878 demonstrates how poorly the country fares on gender issues. The country's GDI is close to the average for all countries in Sub–Saharan Africa, but below that of countries such as Tanzania and Madagascar. Female adolescents in Uganda fare particularly poorly. DHS data from 2011 suggests that dropout from secondary schools is significantly larger for girls than for boys – although primary school completion rate for boys (68 percent) is not very different from that of girls (66 percent)[3], there is a large disparity in their secondary school completion rate for boys (52 percent) and girls (24 percent).

The gender divide in Uganda is evident when examining marriage patterns – median age at marriage for men in 2011 was 22.3 years, while for women it was around 18 years. While only 9 percent of men were married by the age

[2]Source: World Bank database – World Bank website. This is in Purchasing Power Parity (PPP) terms.

[3]From DHS reports

of 18, nearly 49% of women were married by that age. A study by Jain and Kurz, 2007 ranked Uganda ninth among the top 20 hotspot countries for child marriage, while another study in 2013 ranked Uganda as 16th among the 25 countries with the highest rates of early marriages, with 46% and 12% of girls married below the ages of 18 and 15 years respectively (World Vision 2013[4]).

The Constitution of Uganda stipulates 18 years as the minimum legal age for marriage for both boys and girls. Despite this provision, child marriage, still persists among many ethnic and tribal groups. Rubin et al., 2009 discuss reasons why the practice is common in Uganda. First, given that pre–marital pregnancy is viewed as a shameful and stigmatizing event in Uganda, parents are likely to view early marriage among girls to be a way of reducing the chance of pregnancies outside wedlock and of thus protecting family dignity. Second, impoverished households that are large might have an incentive to marry girls off at an early age because of the common custom of bride price under which the household of the groom gives the household of the bride a relatively large amount of money at the time of the wedding (Rubin et al., 2009, Lubaale, 2013).

The Constitution of Uganda stipulates 18 years as the minimum legal age for marriage for both boys and girls. Despite this provision, child marriage, still persists among many ethnic and tribal groups. Rubin et al., 2009 discuss reasons why the practice is common in Uganda. First, given that pre–marital pregnancy is viewed as a shameful and stigmatizing event in Uganda, parents are likely to view early marriage among girls to be a way of reducing the chance of pregnancies outside wedlock and of thus protecting family dignity. Second, impoverished households that are large might have an incentive to marry girls off at an early age because of the common custom of bride price under which

[4]UNICEF Website – Accessed on 31 January 2017

the household of the groom gives the household of the bride a relatively large amount of money at the time of the wedding[5] (Rubin et al., 2009, Lubaale, 2013).

3.3 Data

For this analysis, I use data from the Uganda Demographic and Health Survey (UDHS) conducted in 2001[6]. The DHS are nationally representative surveys that collect information on a wide range of population and health indicators. I focus on the module that was administered to women between the ages of 15 and 49 years, which covered topics such as household characteristics, schooling, labour force participation, fertility, infant and reproductive health, antenatal and post-natal care. This section of the survey contains one of the key variables around which I set up my empirical strategy – the age at which women experienced the onset of menarche. I restrict my sample to all the female respondents who had data available on this variable. Below I describe the different outcomes I examine for the women in my sample. In addition, I also focus on health outcomes for children between the ages of zero and five years who were born to these women.

Since this study is focused on identifying the impacts of women's age at marriage, women who report the age at which they got married provide the sampling frame for my analysis – a total of 5643 women. I drop 315 of these observations since they are missing information on age at menarche, which I use

[5]In a study on Uganda, Bishai and Grossbard, 2010 find that the size of the transfers are around 14 percent of household annual income. See Kaye et al., 2005, Huzayyin and Acsadi, 1976, Dekker and Hoogeveen, 2002 for a detailed discussion on bride price and associated concepts.

[6]Although there are more recent rounds of data available for Uganda, I use this dataset because this is the only DHS data from Uganda that has information on the age of menarche.

as an instrument, or women who report outlier values on this variable. Subsequently, I also exclude 379 women who do not have data on height since this is an important control variable in my empirical specifications. The remaining 4949 women constitute the sample for the woman–level analysis that I conduct. To understand whether the criteria used to identify the woman-level sample lead to a systematically different sample , I conduct two sets of mean–difference analyses: first, I compare women who report marriage age (5643 women) with women who compose the sample for the main analysis (4949 women). This is presented in table A3.1 – the results indicate that there are no differences between the two samples. Second, table A3.2 presents results for the comparison between women who are part of the sample for the analysis (4949 women) and the ones who drop out due to the different criteria described above (315 + 379 = 694 women). While there appear to be no statistically significant differences on most characteristics, I find that the women in the study sample are less likely to be married to men who have multiple wives and more likely to be from the western part of Uganda. Note that these are covariates that I control for in all specifications.

The child–level sample for this analysis consists of the children of the 4949 individuals in the woman–level sample who are five years or below at the time of the DHS survey. The data indicates that 3998 of the women have children in this age group – a total of 5022 children. I conduct the following mean difference tests: first, I compare women who report marriage age (5643 women) with women who have children in the sample for the child–level analysis (3998 women). This is presented in table A3.3 – the results indicate that there are no differences between the two samples. Second, table A3.3 presents results for the comparison between women who are part of the sample for the analysis

(3998 women) and the ones who do not have a child in the child–level analysis (5643 – 3998 = 1645 women). These results are presented in table A3.3. While there appear to be no statistically significant differences on most characteristics, I find that the women in the study sample are less likely to be married to men who have multiple wives – I control for this in all child–level analysis empirical specifications.

3.3.1 Key Variables

The DHS contains data on several outcomes that might be impacted by a woman's marriage age – those pertaining to her educational attainment, labour market participation and health knowledge. To measure literacy, I create a categorical variable that takes a value of one if a woman is fully literate (able to read and write in her native language), and zero otherwise. The DHS enumerators were trained to judge women's literacy levels based on her ability to read sentences printed on cards. I use a continuous variable for educational attainment that equals the highest grade attained by a woman (in years). Woman's labour force participation is a dummy variable equal to one if she reports being part of the labour force at the time of the survey. In examining health practices, I use indicator variables for the use of contraceptives and for the obtainment of Ante Natal Care (ANC) during pregnancy. I also examine effects of marriage age on the age at which the woman gives birth to her first child.

The intergenerational health outcomes that I probe are height (measured in terms of height for age z–score), weight (weight for age z–scores) and BMI (measured in kg/m^2). I construct Z–scores for height and weight based on standard

World Health Organization (WHO) definitions. I also examine children's blood hemoglobin levels using absolute hemoglobin values and also with indicator variables that capture whether children are anemic (below 11 g/dl) or severely anemic (below 7 g/dl)[7].

In seeking to identify the factors that might mediate the relationships I observe between women's marriage age and her later life outcomes, I investigate the role of marriage *quality* and women's decision making power in their post–marriage household. I measure marriage quality using three different measures. I create a variable for spousal education (measured in years of education) which is akin to the educational attainment variable for women. The wellbeing effects stemming from the human capital of the husband and wife would further be reinforced by the positive synergy that is likely to result if the marital relationship were an equitable one (Schultz, 1990, Engle, 1997). While having more educated spouses might be considered to be desirable, an increase in the education of a spouse relative to that of the woman could have a countervailing negative effect. I measure the potential for such an effect using a spousal education gap variable which takes a value that is equal to the difference in the educational attainment of a woman and her husband. Analogously, a high age difference between a woman and her spouse could skew the balance of power in the household in the favor of the husband and to probe whether this is the case, I construct an age gap variable (Basu and Koolwal, 2005, Mahmud et al., 2012).

The DHS asks female respondents a range of questions about her role in decision making on the following topics – own health care, children's health care, large household purchases, daily purchases, visits to family and friends, and items cooked in the household. The survey allowed for several responses

[7] I define these variables in accordance with the WHO guidelines on Anemia detection – Link.

in order to capture different levels of involvement in these decisions[8][9]. I create categorical variables for each decision area that take a value of one if the woman reports making a particular decision individually since this indicates full female autonomy.

3.4 Empirical Strategy

Using Ordinary Least Squares (OLS) to examine the relationship between women's marriage age and an outcome such as educational attainment, would likely produce biased results. These biases could occur due to unobservable factors that might shape both the outcome of interest and the main explanatory variable (marriage age). For example, the traditional beliefs of a woman's natal family could have an impact on how long she stays in school, and also determine when she gets married. Since all such potential confounders cannot be directly observed or measured, endogeneity bias tends to be unavoidable when using the OLS framework. To overcome these issues, I use an Instrumental Variable (IV) strategy, where I treat woman's age at menarche as an instrument for her age at marriage[10].

[8]The responses are as follows: *Respondent alone*, *Husband/Partner alone*, *Respondent and Husband/Partner jointly*, *Someone else individually*, *Someone else and respondent jointly*, *Not Applicable*.

[9]Based on the past literature on the subject, I argue that the recorded responses to these questions are credible proxies for the different dimensions of women's bargaining power – their sense of entitlement and confidence (Kabeer, 1998, Taylor and Pereznieto, 2014), female access to economic resources within the household (Kabeer, 2008) and women's ability to interact/socialize with people outside the household (Kabeer, 2011). In more recent developments, factor analysis has been used to create an index of female empowerment based on the responses to the different bargaining power questions in surveys like the DHS (example – Yount et al., 2016, Cheong et al., 2017). I create a similar measure and check the robustness of the main results with this new variable and find qualitatively similar results.

[10]Age of menarche has also been used as an instrument in many developed country contexts. For example, see Klepinger et al., 1999, Chevalier and Viitanen, 2003, Sabia and Rees, 2009, 2011.

3.4.1 Age of Menarche as an IV

Child marriage is commonly practiced in many developing countries. Since the ability to bear children is an important part of marriage in these contexts (Anderson, 2007), girls are typically married only after they have reached puberty. Thus the age at menarche (first period) tends to be a strong determinant of female age of marriage (Field and Ambrus, 2008a). Since age at menarche is primarily determined by genetic factors, this variable provides the quasi–random variation in age at marriage required to uncover its causal effects on subsequently realized outcomes. I thus use age at menarche as an instrumental variable for the timing of marriage in a Two Stage Least Squares (2SLS) estimation strategy.

The methodological approach that I use requires that the instrumental variable meet two conditions – the inclusion and the exclusion restriction. The inclusion restriction requires that the instrument, or menarche age, be a strong predictor of the potentially endogenous variable, which in my case is women's marriage age. Figure 3.1, which presents the relationship between these two variables, shows that the distribution of marriage age is a parallel but shifted version of that of menarche age, with the peak of the former being to the right of the highest point of the latter. This kind of a relationship would arise if parents married off their daughters shortly after the onset of puberty. In fact, this is consistent with what I find in my dataset – nearly 72 percent of the women in my sample report marrying within three years of the onset of puberty. Figure 3.1 also demonstrates that as age of menarche goes up, age of marriage also goes up, thus showing the tight co–movement of the two measures.

Insert Figure 3.1

I further examine the relationship between the ages of menarche and marriage using regression results. In Table 3.2, I look at the relationship between menarche age and marriage (and other controls that I discuss below), which represents the first stage of the 2SLS estimation which captures the relationship between the instrument (age of menarche) and the potentially endogenous main variable of interest (marriage age). The results indicate that each year of delay in menarche increases marriage age of a woman by around 0.5 years. This relationship is statistically significant at the one percent level and is robust to the addition of a large number of control variables. The F-statistic of the excluded regressor in the first stage is well above the critical value of 10 (the cutoff suggested by Staiger and Stock [1997] for a weak instrument). Given the evidence from figure 3.1 and the first stage results in table 3.2, it seems more likely that menarche age meets the inclusion restriction requirement for a valid instrument.

Insert Table 3.2

Women's age at menarche also needs to meet the exclusion restriction so that it can serve as a valid instrument for marriage age. Under this restriction, the instrument can impact the outcomes of interest through no channels apart from the endogenous variable, but this condition is not directly testable (Bound et al., 1995, Angrist and Krueger, 2001). Although the exclusion restriction is not directly testable, I argue that age of menarche is exogenous as it is biologically determined, which implies that the exclusion restriction plausibly holds in this setup. Having said that, there are ways in which the exclusion restriction can be violated – I discuss them below and also provide the steps I take to account for them.

One potential concern is that the onset of puberty could be shaped by a woman's early life socioeconomic and nutritional conditions, which in turn could also influence her later life outcomes (Freedman et al., 2005). This would make age at puberty, the instrument, endogenous with later life outcomes. In fact, some studies have shown that early life circumstances play an important role in determining menarche age[11] (Berkey et al., 2000, Chowdhury et al., 2000, Ellis, 2004). In contrast, other evidence suggests that genetic composition at birth matters more for puberty onset than post–birth environmental factors (Jahanfar et al., 2013, Sørensen et al., 2013, Adair, 2001). Along these lines, a recent study (Mpora et al., 2014) finds that early life adverse events do not have a significant effect on age at menarche in Uganda, the context that I examine in this study. Given the debate that exists on this topic, it is unclear whether age of menarche can truly be considered to be an exogenous variable. One way in which I could overcome potential endogeneity in the instrument is by including measures for women's early life conditions in my estimation models.

While the ideal way to control for a woman's childhood nutritional status would be to use information from that time, the data I use does not contain such details. Therefore, in order to account for the role of childhood nutrition in determining menarche age (Ellis, 2004, Victora et al., 2008), I use woman's adult height as a control variable. This approach is predicated on the intuition that a woman's adult height reflects her childhood height, which itself corresponds closely to childhood nutritional status (Martorell and Habicht, 1986a, Martorell,

[11] This is consistent with literature that shows the negative effects of shocks in the prenatal and perinatal period (Barker, 1995, Almond and Currie, 2011b) and in early life (Dercon and Porter, 2014, Almond, 2006, Fogel, 1993, 1990).

1993). Adult height can thus be used to proxy for the different inputs experienced by women during childhood. In addition, research indicates that people with lower stature in infancy and childhood are more likely to have lower adult height (Sørensen et al., 1999, Eide et al., 2005, Adair, 2007, Currie and Vogl, 2013). As adult height is strongly correlated with childhood size and nutritional inputs in childhood, including it in my empirical model would control for the effect of childhood stature on menarche.

In addition, I use birth year fixed effects to control for the effect that events in infancy can have on long term outcomes. I also include district fixed effects and cluster altitude (in meters) in my specifications to account for the potential consequences of geographical conditions (such as temperature and altitude) and other time–invariant district level factors on the age of puberty onset (Kapoor and Kapoor, 1986, Saar et al., 1988).

Another concern that arises when using reported age at menarche stems from potential recall bias. Given that a considerable amount of time is likely to have passed from the date when a women attained menarche (recall that survey respondents could be a maximum of 49 years of age at the time of the survey), one might question the ability of women to accurately report their age at puberty onset. To get a sense of how dependable the reports of menarche are in the DHS dataset that I use, I compare the mean menarche age in my sample (14.4 years) with information available from other parts of Africa. I find that the mean in my sample is broadly consistent with that of other studies that examine comparable settings – such as Padez, 2003 in Mozambique (13.2 years), Leenstra et al., 2005 in Kenya (15.8 years), Zegeye et al., 2009 in Ethiopia (15.7 years) and Adebara and Ijaiya, 2013 in Nigeria (13.2 years). It is also worth pointing out that in the cultural context of many developing countries (such as Uganda), onset of menarche is a major event in a woman's life and hence respondents could be expected to remember its timing with a fair degree of accuracy[12] (Leenstra et al., 2005, Ellis, 2004).

3.4.2 Identification Strategy

As discussed above, I examine outcomes for women and for their children. When examining women's outcomes, I use the following specification:

$$MarriageAge_j = \alpha_0 + \alpha_1 MenarcheAge_j + \alpha_2 Controls_j + \eta_j^1 Y_{ij} = \delta_0 + \delta_1 MarriageAge_j + \delta_2 Controls_j + \eta_j^2 \tag{3.1}$$

where Y_j is the outcome variable for woman j, $MarriageAge_j$ is the age at marriage of woman j, $MenarcheAge_j$ is the age at which a woman hits puberty and $Controls_j$ is a vector of factors at the individual, household and community level that could potentially shape the outcome of interest including household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and residence in an urban area. In addition, I also include woman's height, and fixed effects for both district and birth year to account for potential non–genetic factors associated with age of menarche. I cluster the standard errors at the district level.

[12] A related concern is that women might recall the age at which they reached puberty using the year in which they got married as a reference point. Since marriage is another major social event in one's life, its recall is also less likely to be fraught with measurement error. Given that women in Uganda tend to get married in and around the time of puberty onset (as discussed earlier), this kind of a connection between the two events could only improve the recall of menarche onset, thus plausibly improving the accuracy of the data.

The specification that I use to analyze child level outcomes is similar, but now each observation pertains to a child born to a woman in the main sample. I use the following approach:

$$MarriageAge_j = \alpha_0^c + \alpha_1^c MenarcheAge_j + \alpha_2^c Controls_{ij} + \zeta_{ij}^1 Y_{ij} = \delta_0^c + \delta_1^c MarriageAge_j + \delta_2^c Controls_{ij} + \zeta_{ij}^2 \tag{3.2}$$

In this case, the outcomes would be for child i born to woman j, $MarriageAge_j$ is the age at marriage of the mother of child i, $Controls_j$ include all the control variables discussed for the mother–level specification as well as age of the child, mother's age at the child's birth, child gender and birth order of the child. Again, I cluster the standard errors at the district level.

Even though most of the outcomes that I examine in this research are binary variables, I use LPM for my main analysis for ease of coefficient interpretation. However, I also estimate probit models to demonstrate the robustness of the main estimates[13].

[13]LPM suffers from the criticism that the error term in the specification is not independent of the covariates in the model (unless there is just a single binary covariate). Since we have more than one covariate, this becomes important. Another drawback of the LPM is that the predicted values of the dependent variable can sometimes be outside the zero to one (feasible values) range. The typical response to these criticisms is that the purpose of the LPM is not to make predictions for the entire support of the covariates, but rather for a subset of the support. Additionally, LPM has a constant marginal effect that might be preferable in a variety of circumstances. In the same vein, Probit and Logit models have their own pros and cons. While they are both non–linear models of binary choice affording more flexibility, they impose strong assumptions on the error term of the structural model. It is hard to check if these assumptions are the right ones for the data provided, unless there is some prior theory that supports these assumptions.

3.5 Results

3.5.1 Impact on Woman's Outcomes

I begin by examining the consequences of women's marriage age on her educational attainment. First in Table 3.3 (columns 1–4), I examine this relationship with an OLS approach and find that marriage age has a positive impact on number of years of education. In other words, women who marry at a younger age appear to have lower educational attainment. The OLS estimates (columns 1–4) are likely to suffer from endogeneity bias, which can be accounted for with an instrumental variable approach. In the IV results which I present in columns 5–8, the coefficients on marriage age are positive and the magnitudes do not change much when I include additional sets of control variables in the models. These findings point to the potentially large gains in female education that can be realized by delaying marriage.

Insert Table 3.3

In table 3.3, the coefficient on marriage age in the IV specification with all the control variables (column 8) is larger in magnitude than the corresponding OLS coefficient (column 4) which suggests that OLS underestimates the effect of marriage age. I observe similar trends for all the other outcome variables that I examine. Such a situation, where the IV estimates are greater than the OLS estimates, could arise due to omitted variable bias. For example, families in Uganda might be making decisions regarding their daughter's schooling and marriage with the goal of maximizing the bride price that they can obtain from the groom's family at the time of the wedding – which I do not observe with the available data. Given that bride price is primarily paid for the fruits of a woman's labour and her reproductive capabilities (Anderson, 2007), girls' parents might try to increase the amount they receive by keeping girls in school longer and thus enhancing girls' future labour market returns. At the same time since bride price also depends on the virginity of the bride, in relatively unsafe areas, parents might want to get their girls married sooner rather than later. The interaction effect of the desire of parents to keep their girls in school longer but to also marry them off younger would attenuate the OLS estimates[14].

Insert Table 3.4

Table 3.4 explores the effect of marriage age on other woman level outcomes. I find that later marriage among women enhances the likelihood of being literate and of participating in the labor market. Specifically, a one year delay in marriage increases the probability of being literate and of working by 10 and eight percentage points respectively. Subsequent columns show that there also are positive impacts of marriage age on different health behavior outcomes – when women marry later, they are more likely to use contraception and obtain antenatal care when pregnant. I find evidence for delays in childbearing due to later marriage, which indicates that policy interventions to encourage delayed marriage could be crucial for alleviating the high maternal health burdens borne by young mothers (O'Flaherty et al., 2015).

[14]The IV estimate being greater than the OLS is also consistent with the theory of positive assortative matching in the marriage market Becker, 1991, which I find evidence for in my data. Positive matching implies that higher quality grooms would get matched to women with more desirable traits – which in this case would imply higher education (stay in school), and virginity (get married sooner).

3.5.2 Intergenerational Impacts

I now explore whether the timing of a woman's marriage shapes her children's health outcomes. Table 3.5 shows that there are no effects of marriage age on standardized height and weight measures of children, but the coefficients are signed as expected with older age at marriage leading to healthier children. I do, however, find positive and statistically significant effects of later marriage on child BMI and hemoglobin level. In other words, when mothers delay marriage by one year, their children in the future are likely to have better BMI and have higher hemoglobin levels. Other results in table 3.5 indicate that later marriage reduces the chances that a woman's child will be anemic (4 percentage points), but the impact on the likelihood of being severe anemic is small (0.2 percentage points) and insignificant.

Insert Table 3.5

3.6 Mechanisms

The results in the previous section present a unified narrative – delayed marriage brings about better later life outcomes for women, as well as positive effects for child health. Since I find that later marriage enhances women's educational levels, many of the other benefits that I find for these women could stem directly from the higher education that they obtain. Other factors that are shaped by later marriage (and that could also be influenced by education) could lead to improved later life outcomes. Here I examine the role played by women's bargaining power and the nature of their marital relationship.

Studies show that an increase in female autonomy could raise the wellbeing of women and also enhance their ability to allocate more resources towards their children (Ashraf et al., 2010, Aslam and Kingdon, 2010, Doss, 2013). Women with more agency might have more autonomy to make decisions regarding contraceptive use and might thus have lower fertility rates (Beegle et al., 2001). Due to the quantity–quality tradeoffs in children (discussed first by Becker and Lewis, 1973), lower fertility is likely to enhance the quality of children, potentially due to higher investments per child (Barber and Gertler, 2010, Carneiro et al., 2013, Björkman Nyqvist and Jayachandran, 2017).

Recall that I measure women's decision making power by observing whether she self–reports being solely responsible for making decisions regarding different aspects of the household. In table 3.6 (columns 1–5), I find that later marriage leads to increase in the likelihood of women being the sole decision maker on every measured decision category – child health (4 p.p.), own health (5 p.p.), large purchases (4 p.p.), family visits (3 p.p.) and cooking food (5 p.p.). This is consistent with the evidence from other recent analyses that find that women who marry later in life benefit from post–marital economic empowerment, have higher decision making power and enjoy more equitable gender relations (Yount et al., 2018, Crandall et al., 2016).

Insert Table 3.6

Next, I examine the potential role of spousal characteristics in mediating the observed effects of later marriage. Previous studies have found that more educated women are likely to marry higher quality spouses (Fafchamps and Quisumbing, 2005, Abramitzky et al., 2011), with quality being defined on many dimensions such as education and income. I verify whether this is the case for the women in my sample who marry later in and who I find are also likely to be more educated. Results in Table 3.6 suggest that women who marry later in life are matched with husbands who have higher education. Other measures of relative marriage market match quality are the differences in educational levels and age between the spouses – lower the difference,the more equitable the marriage is likely to be. The coefficients on marriage age for both outcomes are negative and significant, thus demonstrating that the spousal educational and age gap falls as marriage age rises. These results indicate that older brides experience improved marriage market outcomes.

3.7 Robustness Checks

I conduct several robustness checks to demonstrate that the results from my main analysis are not sensitive to model variations, definitional adjustments and other changes. First, I estimate probit models for the outcome variables that are binary and examine whether the results are consistent with those from the LPM model that I use for my main analysis. Results in Table 3.7 indicate that the sign and significance of the marriage age coefficients in the probit models are consistent with those in the main results. In addition, the marginal effects from the Probit model are comparable in magnitude to the estimates identified with the LPM specification.

Second, I re—estimate the specifications for all the outcomes that I examine with alternative measures for early marriage. While I use a continuous measure for marriage age in the main analysis, I now choose three different cutoffs

(18, 16 and 14 years) to create binary variables indicating early marriage – these measures consider women marrying at ages below the thresholds to have married early. A cutoff of 18 years is almost universally accepted as an appropriate minimum age for marriage and hence is a useful threshold to examine. In the Ugandan context, using cutoffs for 16 and 14 years also makes sense as almost 34% and 18% of the study sample marry under these respective ages. The results in table 3.8 show that my results are robust to these alternative definitions of marriage age (with one exception). I find that women identified as having married early by these categorical variables have poorer later life outcomes and less children.

Insert Table 3.8

Third, I conduct a check to predict what would happen to the identified impact estimates if the exclusion restriction were to be violated. In this case the exclusion restriction requires that age of menarche affect later life outcomes only through its impact on marriage age, and not through any other variables. Through the check, I seek to understand whether the results would hold if there was a non–zero direct relationship between the instrument and the outcomes of interest. I employ the Union of Confidence Intervals (UCI) procedure outlined in Conley et al., 2012, where the authors relax the complete exogeneity assumption made in an IV setup[15]. This method requires that the researcher specify a value for the (assumed) direct relationship between the instrument and the outcome (referred to by the authors as γ). The γ term can be thought of as a measure of the degree to which the exclusion restriction is violated. This procedure calculates the confidence interval for the coefficient of interest (marriage

[15]For a detailed discussion on this, please look at Conley et al., 2012.

age) for a specified value of γ. If this confidence interval contains the value zero, it indicates that the coefficient is statistically indistinguishable from zero.

In assuming different values of γ, I start with zero (complete exogeneity) and gradually go up to 0.25. The cutoff of 0.25 is arbitrary, but is a fairly high number – the correlation between menarche age and any outcome of interest is not expected to be this high due to the extensive set of control variables in the estimation models. Upon generating the relevant confidence intervals, I find that the study results persist for most variables[16].

Since this analysis explores the impact of marriage age on a large number of outcomes, I also calculate multiple hypothesis testing adjusted p–values to understand whether the results I find are spuriously significant. Given that a large number of my point estimates are significant and they fit into a consistent narrative, it is unlikely that such corrections would largely change my conclusions. I present the results adjusted for multiple testing (with the Romano–Wolf procedure procedure (Romano and Wolf, 2005a,b, 2016) in Table 3.9 – only one outcome (Child BMI) goes from being significant to not significant after adjusting for multiple hypothesis testing. The coefficients that were not significant to begin with remain so.

As a final check, I explore whether my results might be impacted by omitted variable bias. Although I control for a large number of factors in my specifications, there is a possibility of there being other factors that affect the outcome

[16] Although this method provides a technique to test the sensitivity of the results to violation of the exogeneity assumption, Conley et al., 2012 state that one of the caveats of this technique is that it might give a wide confidence interval, which might not be very informative. However, the wide confidence intervals actually makes the test harder to pass– if the confidence intervals are larger, then they are more likely to include the null result. Therefore, if the results hold for fairly large values of γ (which I show), then they could be interpreted as being extremely robust.

which I am unable to account for[17]. to check for this, I employ the bounding exercise undertaken in Oster, 2014[18]. Under this robustness check, one case would be to assume that the bias due to unobservables is of the same size as the bias due to observables – this is an extreme assumption because it implies that the unexplained part of the regression is as large as the explained part of the regression, an implausible contingency since I control for a large number of individual, household and community factors in all the specifications. I find that even under such a restrictive assumption, the impact of marriage age retains the direction of the impacts identified in the main analysis and does not move towards zero for 9 out of 12 outcomes[19].

3.8 Conclusion

Using a nationally representative dataset from Uganda and an IV estimation strategy, I provide causal evidence for the effects of marriage age on a variety of later life outcomes. Early marriage reduces women's educational attainment, literacy and labour force participation, and leads to poor health behaviors. I also find that the lower the age at which women get married worse their outcomes are in terms of contraceptive use, age at first birth and usage of ante natal care during pregnancy. Furthermore, I find that marriage age shapes intergenerational health outcomes – I find strong negative effects of early marriage on child hemoglobin levels, anemia and BMI. While the impacts on child height

[17]For example, we do not observe the educational/health infrastructure that women experienced when they were children, which would have an impact on their later life outcomes.

[18]Under this method, she extends the theoretical framework proposed by Altonji et al., 2005 to connect bias on unobservables to coefficient stability. She uses the movement of the R–squared value with and without the controls along with the potential size of the bias of the unobservables to estimate lower bounds on the reduced form estimate.

[19]Results available from the author on request

and weight are not statistically significant, the direction of effect implies that earlier marriage leads to worse child health status. In examining the potential mechanisms through which the observed effects might be mediated, I detect strong positive impacts of later marriage on women's decision making power and status, and evidence of positive assortative marriage market matching (that is women who marry later are more likely to attract higher quality spouses). These factors, along with the higher educational levels stemming from delayed marriage among women, might be responsible for the many positive later life outcomes for women and the improvements to child health that I find evidence for in my analysis.

Further, this paper adds to the literature by bringing forth evidence in a context where bride price is commonly practiced. This is a valuable contribution since the bulk of the existing economic literature on the topic has focused on the effects of marriage age in developing country contexts where the practice of dowry (payments from the bride's family to the groom's family) is prevalent. For example, Sekhri and Debnath, 2014 and Chari et al., 2017 find that marrying at an older age has positive effects on a number of later life outcomes for women and their children in India, a setting where dowry is commonly practiced. Similar results have been found in other dowry contexts – Bangladesh (Field and Ambrus, 2008a), Egypt (Yount et al., 2018) and Kenya (Hicks and Hicks, 2015).

Additionally, my findings are consistent with those of the aforementioned studies – earlier marriage has a negative impact on a variety of later life socio–economic outcomes. The findings are similar in other regions where neither dowry nor bride price is widely practiced (such as USA. Studies have found detrimental effects of child marriage on maternal education and employment, as well as on maternal and child health outcomes (Overpeck et al., 1998, Hotz et al., 1997, Hunt, 2003, Dahl, 2010). The fact that my results qualitatively match those found in areas with different marriage customs suggests that the negative effects of early marriage is a global phenomenon and is independent of the specific features of the marriage market. Having said that, the exact mechanisms through which this effect operates may be more context–specific.

Although there has been considerable global reduction in the prevalence of child marriage over the past couple of decades, it still remains a major concern today. According to UNICEF, nearly 650 million girls marry before the age of 18 years (UNICEF, 2018). Given the persistence of this problem, ending child marriage has been included as part of the UN's Sustainable Development Goals (goal 5.3). Child marriage in Uganda in recent years mirrors broad global trends. The country has experienced a decline in marriage age – from 52.8 percent in 2001 to 36.5 percent in 2011, but absolute levels of child marriage in the country remain high (Male and Wodon, 2016, Wodon et al., 2017). Given that Uganda's population is extremely young (6 in 10 people are below 18 years of age (Sebudde et al., 2017) and that the practice of child marriage remains common, a large section of the country's population is likely to continue to experience the far reaching harms of early marriage documented in this research. It is, thus, crucial to design policies to limit this practice in Uganda and around the world.

In policy terms, developing country governments can reap potentially large (and long–term) gains by taking steps to restrict child marriage practices. Various strategies for delaying marriage among young girls have been studied. One approach has been to encourage girls to stay in school by providing school vouchers or stipends. Evaluations of such programs in Colombia, Bangladesh and Kenya (Angrist et al., 2006, Duflo et al., 2015, Hahn et al., 2018, Hong and Sarr, 2012) show promising results. In a unique cash transfer experiment in Malawi, Baird et al., 2011 find that the unconditional transfers outperform the conditional ones in their effectiveness to impact early marriage and child bearing outcomes. Another strategy is to expose girls to material and information that enhance female empowerment. For example, girls can be provided with vocational training and with information on sex, reproduction and marriage. Such policies have been tested in Uganda (Bandiera et al., 2014), Tanzania (Buehren et al., 2015) and Bangladesh (Buchmann et al., 2017) with varying degrees of success. Bandiera et al., 2014 find that over the two year study span in Uganda there were pronounced declines in early marriage (58%), and teen pregnancy (26%).

In 2015, Uganda instituted a National Strategy on Child Marriage (NSCM), under which it aims to end child marriage and teenage pregnancy. Since child marriage is deeply embedded in Ugandan society, the NSCM intends to use a multi–pronged approach to achieve its objective – improvements to the country's legal and policy framework for protecting children, expansion of education, empowerment of boys and girls through the provision of information and critical life skills, and changes to the mindset and beliefs of different communities/tribes regarding this practice. While Uganda's program appears to be a holistic approach for ending child marriage, its success would largely depend on the way in which the plan is implemented. The results of this study suggest that if Uganda's strategy proves to be effective in reducing child marriages, the country is likely to have a more educated and empowered female population, and healthier children in the future.

3.9 Figures

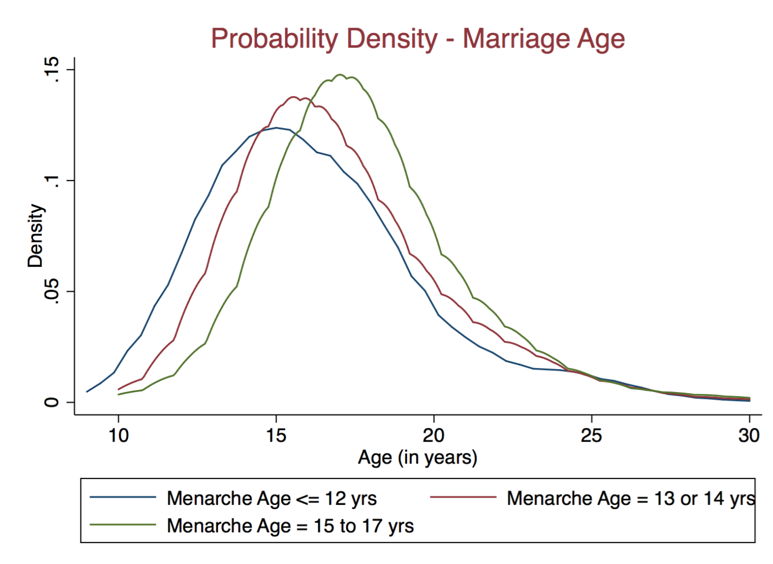

Figure 3.1: Marriage age distributions.
This clearly shows that marriage age increases as the age at menarche rises.

3.10 Tables

Table 3.1: Mean Comparison Across Marriage Ages

	Less than 14	14–16 years	16–18 years	Above 18 years	Full Sample
Marriage Age	13.18	15.54	17.47	21.42	17.37
	(1.024)	(0.499)	(0.499)	(2.915)	(3.419)
Menarche Age	13.71	14.27	14.64	14.70	14.40
	(1.301)	(1.296)	(1.462)	(1.508)	(1.453)
Multiple Wives	0.305	0.245	0.237	0.255	0.257
	(0.461)	(0.430)	(0.426)	(0.436)	(0.437)
Wealth Index	-0.0155	-0.00531	0.0263	-0.0127	-0.000891
	(1.029)	(1.016)	(1.017)	(0.985)	(1.009)
HH size	5.915	5.650	5.637	5.493	5.647
	(3.099)	(2.782)	(2.979)	(2.968)	(2.951)
Urban	0.231	0.250	0.268	0.374	0.288
	(0.422)	(0.433)	(0.443)	(0.484)	(0.453)
Telephone	0.0158	0.0247	0.0359	0.0808	0.0427
	(0.125)	(0.155)	(0.186)	(0.273)	(0.202)
Altitude (in mts.)	1278.8	1287.7	1305.4	1345.2	1307.8
	(295.5)	(308.1)	(291.7)	(340.0)	(312.5)
East	0.306	0.298	0.261	0.201	0.261
	(0.461)	(0.457)	(0.440)	(0.401)	(0.439)
North	0.158	0.170	0.125	0.123	0.142
	(0.365)	(0.376)	(0.331)	(0.328)	(0.349)
West	0.226	0.237	0.285	0.317	0.271
	(0.418)	(0.425)	(0.452)	(0.465)	(0.445)
Brick House	0.300	0.306	0.333	0.406	0.341
	(0.458)	(0.461)	(0.471)	(0.491)	(0.474)

This table provides summary statistics of the variables used in the analysis for the different age ranges of marriage age in the sample. The last column provides the summary statistics for the full sample.

Table 3.2: First Stage Regressions – Dependent variable is Age at First Marriage

	FS 1	FS 2	FS 3	FS 4
Menarche Age	0.49***	0.48***	0.47***	0.45***
S.E. (coef)	0.03	0.04	0.03	0.03
Total Obs.	4930	4930	4930	4930
Controls	No	Yes	Yes	Yes
Birth Year FE	No	No	Yes	Yes
District Dummy	No	No	No	Yes

The outcome variable is women's marriage age. The standard errors are robust and clustered at the district level. The control variables include woman's height, woman's age at first sex (intercourse), cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area.

Table 3.3: Impact of Marriage Age on Highest Grade Attained

	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Marriage Age	0.25***	0.26***	0.25***	0.26***	0.74***	0.75***	0.73***	0.75***
S.E. (coef)	0.02	0.02	0.02	0.02	0.08	0.08	0.08	0.08
Total Obs.	4912	4912	4912	4912	4912	4912	4912	4912
Controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Birth Year FE	No	No	Yes	Yes	No	No	Yes	Yes
District Dummy	No	No	No	Yes	No	No	Yes	Yes

The outcome variable is women's highest grade attained. The coefficients are from the second stage of a 2SLS IV estimation. The standard errors are robust and clustered at the district level. The control variables include woman's height, woman's age at first sex (intercourse), cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area.

Table 3.4: Impact of Marriage Age on Women's Later Life Outcomes

	Literacy	Labor	Contraception	Age FB	ANC Usage
Marriage Age	0.10***	0.08***	0.02*	1.01***	0.07
S.E. (coef)	0.01	0.05	0.01	0.01	0.07
Total Obs.	4899	4920	4925	4916	4522

Following are the definitions of the outcome variables: Literacy (=1 if literate), Labor (=1 if part of the labour force), Contraception (=1 if used contraceptive), Age FB (woman's age at first birth), ANC Usage (=1 if reported using ante natal care). The coefficients are from the second stage of a 2SLS IV estimation. The standard errors are robust and clustered at the district level. The control variables include woman's height, woman's age at first sex (intercourse), cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area. All specifications also include birth year and district fixed effects.

Table 3.5: Impact of Marriage Age on Children's Health Outcomes

	HFA Z–Score	WFA Z–Score	BMI Z–Score	Hemoglobin	Anemic	Severely Anemic
Marriage Age	0.02	0.001	0.11*	0.18***	-0.04***	0.002
S.E. (coef)	0.06	0.05	0.06	0.05	0.01	0.01
Total Obs.	4997	4982	4990	4956	4955	4958

Following are the definitions of the outcome variables: HFA (Height for Age Z–score), WFA (Weight for Age Z–Score), BMI Z–score, Hemoglobin levels (g/dl), Anemic (=1 if hemoglobin below 11 g/dl) and Severely Anemic (=1 if hemoglobin below 7 g/dl). The coefficients are from the second stage of a 2SLS IV estimation. The standard errors are robust and clustered at the district level. The control variables include age of the child, mother's age at the child's birth, child gender and birth order of the child, woman's age at first sex (intercourse), cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area. All specifications also include birth year and district fixed effects.

Table 3.6: Potential Mechanisms

	Child Health	Own Health	Large Purchases	Visit Family	Cook Food	Spouse Edu	Spouse Edu Gap	Spouse Age Gap
Marriage Age	0.04***	0.05***	0.04***	0.03***	0.05***	0.39***	-0.35***	0.35*
S.E. (coef)	0.01	0.01	0.01	0.01	0.01	0.10	0.09	0.19
Total Obs.	4912	4912	4912	4912	4912	4695	4695	4695

The first five columns take a value of one if the woman reports making decisions about these items. Spouse Edu refers to the highest grade attained by the husband, Spouse Edu gap refers to the difference in highest grade attained between the woman and her husband, and Spouse Age Gap refers to the age gap between the woman and her husband. The coefficients are from the second stage of a 2SLS IV estimation. The standard errors are robust and clustered at the district level. The control variables include woman's height, woman's age at first sex (intercourse), cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area. All specifications also include birth year and district fixed effects.

Table 3.7: Robustness Check – Probit model

	Literacy	Labor	Contraception	ANC Usage	Anemic	Severely Anemic
Marriage Age	0.13***	0.12***	0.08*	0.11	-0.07***	0.001
S.E. (coef)	0.06	0.05	0.06	0.05	0.01	0.01
Total Obs.	4901	4921	4924	4524	4955	4955
Marginal Effect	0.08***	0.09***	0.03*	0.08	-0.04***	0.0005

The definitions of the variables are as explained in the previous tables. Each coefficient comes from a probit model where the outcome variable is regressed on the woman's marriage age and a set of controls (discussed earlier). The standard errors are robust and clustered at the district level. The control variables include woman's height, woman's age at first sex (intercourse), cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area. All specifications include birth year and district fixed effects. The marginal effects are calculated at the mean value of all the covariates in the model.

Table 3.8: Robustness Checks – Different Definitions of Marriage Age

Outcome	Under 18 Years	Under 16 Years	Under 14 Years
Education	-0.76***	-0.47***	-0.58***
Labor market participation	-0.07	-0.04	-0.05
Literacy	-0.08***	-0.06***	-0.07***
Contraception use	-0.09***	-0.12***	-0.12***
Child hemoglobin	-0.11***	-0.11***	-0.12***
Child– anemic	0.06***	0.06***	0.07***
Child– severely anemic	0.04	0.04	0.03

The definitions of the variables are the same as that described earlier, and so are the control variables. Each column describes the results when the marriage age is defined to be a categorical variable taking a value of 1 if the woman got married below the age of 18, 16 and 14 years. Each entry is the coefficient of this new marriage age variable when a particular outcome of interest is chosen in the 2SLS framework used for the main set of results. These results provide a robustness check of the main results to a change in the definition of the marriage age variable.

Table 3.9: Multiple Hypothesis Testing– Romano Wolf

Outcome	Regular P-value	Adjusted P-value
Education	***	***
Labor market participation	**	*
Contraception use	*	*
Time to first child	***	***
Age at first birth	***	***
Antenatal Checkup	NS	NS
Child BMI	*	NS
Child hemoglobin	***	**
Child– anemic	***	***
Child– severely anemic	NS	NS
Decision– Child Health	***	**
Decision– Own Health	***	***
Decision– Large Purchases	***	**
Decision– Visit Family	***	**
Decision– Cook Food	***	**

The table shows the significance level of the marriage age variable in regressions where the mentioned variables are the dependent variable. The regular p–value refers to the level of significance in the main results, whereas the adjusted p–value refers to the p–values obtained from the Romano–Wolf procedure. ***, ** and * represent significant at one, five and ten percent respectively. NS symbolizes that the coefficient on the marriage age variable is not statistically significant.

APPENDIX A

APPENDIX – CHAPTER 1

Under the global approach, an RD setup can be estimated using an Instrumental Variable (IV) strategy, where the allocation rule on either side of the cutoff provides the IV. The programme assignment rule used by DPEP, that provide treatment to districts depending on whether their female literacy levels were more or less than the national average, allows me to construct this IV. In this setup, whether a district's female literacy level (in 1991) was above or below this cutoff (39.2 percent) is the instrument. This instrument is is highly predictive of whether or not a district receives the DPEP programme ($DPEP_d$). I create a categorical variable ($BelowAvg_d$) that takes a value of one if the district to which the individual belongs lies below the literacy cutoff (39.2 percent), and takes a value of zero otherwise. The first and second stage equations of this Two Stage Least Squares (2SLS) approach can be written as:

$$DPEP_d = \alpha_1 * BelowAvg_d + \alpha_2 * BelowAvg_d * (DFLR - 39.2) +$$
$$\alpha_3 * BelowAvg_d * (DFLR - 39.2)^2 + \gamma * X_{idt} + \nu_{idt} \quad \ldots \quad [FIRSTSTAGE]$$
$$Y_{idt} = \beta_1 * DPEP_d + \beta_2 * BelowAvg_d * (DFLR - 39.2) +$$
$$\beta_3 * BelowAvg_d * (DFLR - 39.2)^2 + \delta * X_{idt} + \epsilon_{idt} \quad \ldots \quad [SECONDSTAGE]$$

To be a valid instrument for programme participation, the categorical variable ($BelowAvg_d$) needs to satisfy two conditions. The inclusion restriction requires that the potentially endogenous independent variable of interest ($DPEP_d$) be correlated with the instrument ($BelowAvg_d$). In other words, the instrument should be a strong predictor of programme participation. This can be directly tested and I present these results later in the paper. The second condition is the exclusion restriction under which the instrument ($BelowAvg_d$) should impact the outcome only through the instrumented variable ($DPEP_d$), and not through other variables. The exclusion restriction is not directly testable, but I argue that it is likely to be satisfied in this setting. In my knowledge, there were no other government programme at that time (or since) that were allocated based on the allocation rule of DPEP. Given that there were no discontinuities in the provision of other government schemes before DPEP, it is apriori unlikely that there would be any discontinuities in outcomes around the female literacy cut-off chosen for DPEP. In addition, through some falsification tests I show that there were no discontinuities in variables that should be unaffected by the programme. Hence this instrument ($BelowAvg_d$) is unlikely to be correlated with any other covariates around this cutoff. I provide a more detailed discussion in the results section.

APPENDIX B

APPENDIX – CHAPTER 2

B.1 Item Response Theory

The test scores used in this paper are constructed using Item Response Theory (IRT). IRT is still an uncommon measure in the education economics literature, apart from a few exceptions (Singh, 2017, Das and Zajonc, 2010). It is, however, used in evaluating results from large–scale tests, such as the Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMMS), and Graduate Record Examination (GRE).

The main principle of IRT is to differentiate between the latent ability of any given student to answer a question correctly and the actual response given. This is done by three different parameters for any given item: the difficulty, discrimination, and the pseudo–guessing parameters.

The Item Response Function (IRF) links the latent ability to the probability of success in that item for any given respondent. Following Singh, 2017; and Das and Zajonc, 2010, we use the three-parameter (3PL) logistic model introduced by Birnbaum, 1968. Given the probability of a correct response X_ig=1 for a given item g, and given ability level , the probability of successful response is:

$$P_g\left(X_{ig}=1|\theta\right)=c_g+\frac{1-c_g}{1+\exp\left[1.7a_g\left(\theta_i-b_g\right)\right]}| \tag{B.1}$$

where b_g is the difficulty parameter, a_g is the discrimination parameter, and c_g is the pseudo–guessing parameter. The difficulty parameter measures the overall difficulty of the item; the discrimination parameter tells how well a given item can differentiate between different levels of ability. Finally, the pseudo–guessing parameter tells how much success in a given item is random and, thus, unrelated to the respondents ability. Setting the pseudo–guessing parameter to zero will yield a two–parameter model (2PL), which we have used in the cases where the maximum likelihood function of the 3PL–model was not converging. We argue this is not an issue, since the test scores that we were able to estimate with the 2PL and 3PL models are very strongly correlated (close to 99%).

For comparing the levels of the test scores between the two countries, we construct the IRT scores from the joint distribution of the scores of the two countries. The advantage of doing this is that the parameters of IRT are estimated jointly for the common items, which renders the scores comparable. For all the regression analysis, we employ IRT scores that were estimated separately for each country, as we estimate country–specific regression models.

B.1.1 Test score comparability across time and space

The test scores in the second grade were administered in Senegal in 1995–6 and in Madagascar in 1997–8 for both French and math. During those school years, there were two tests for both French and math, one at the beginning of the second grade and one at the end of the second grade, which we call pretest and posttest, respectively. During 2012, French and math tests were administered in both Senegal and Madagascar. These tests were different from the tests administered by PASEC for the second graders.

Table A.1 below reports which data sets are similar across space and time. In the regressions we use IRT scores that are calculated for each country; hence, we do not exploit the comparability in the regressions, given that we run regressions separately for each country. In comparing the difference in performances across Senegal and Madagascar, we use the property that the tests are either fully or partially the same (Figure 1). Table A.1 below explains what the similarities are in different tests across time and space. Notice that tests administered in the second grade are different from those administered in 2012.

		Madagascar (MD)		Senegal (SN)	
		Math	**French**	**Math**	**French**
Children	**Second grade pretest**	No overlap with other tests	No overlap with other tests	Partially same as posttest in SN and posttest in MD.	Same as post-test in SN and posttest in MD.
Children	**Second grade posttest**	Partially same as pre-test in SN, same as posttest in SN	Same as posttest and pretests in SN	Partially same as pretest in SN and as posttest in MD	Same as pretest in SN, and posttest in MD
Adults	**2012**	Partially same in test as 2012 SN	Partially same test as 2012 SN	Partially same test as 2012 MD	Partially same test as in 2012 MD

Figure B.1: Comparison of tests questions

The rows indicate the timing of the test and the columns the country of the test and whether the test is math or French. Each cell includes an explanation of whether any test of the same discipline in the other country was fully the same (exactly the same test questions), partially same (some overlap in the test questions), or completely different (such that no overlap in test questions between countries

Table A2.1: Summary Statistics

SENEGAL	Obs	Mean	Std. Dev.	Min	Max
Highest Grade in 2012	447	8.97	3.82	0.00	15.00
French 2nd grade (pre)	447	-0.06	0.84	-1.47	1.89
French 2nd grade (post)	447	-0.08	0.87	-2.14	2.19
Math 2nd grade (pre)	447	-0.08	0.90	-1.69	2.60
Math 2nd grade (post)	447	-0.07	0.92	-2.59	2.22
Math and French 2nd grade (post)	447	-0.08	0.88	-2.65	2.16
Math and French 2nd grade (pre)	447	-0.05	0.90	-2.02	2.52
2012 Math score	381	0.46	1.43	-3.24	2.90
2012 French score	381	0.46	0.79	-0.86	1.91
2012 Math–French score	381	0.24	0.83	-2.48	1.30
Height in 2012	447	171.88	8.81	149.00	195.00
Female	447	0.42	0.49	0.00	1.00
Age 2012	447	23.80	2.04	16.00	29.00
Mother Education (Dummy)	447	0.09	0.29	0.00	1.00
Father Education (Dummy)	447	0.17	0.38	0.00	1.00
Assets 2nd grade	447	-0.29	0.89	-1.40	1.49
MADAGASCAR	Obs	Mean	Std. Dev.	Min	Max
Highest grade in 2012	333	10.04	3.22	1.00	15.00
French 2nd grade (pre)	333	0.10	1.00	-2.11	2.70
French 2nd grade (post)	333	-0.09	0.99	-2.36	2.51
Math 2nd grade (pre)	333	0.06	0.96	-2.79	1.80
Math 2nd grade (post)	333	0.01	0.89	-2.42	2.15
Math and French 2nd grade (post)	333	-0.04	0.94	-2.43	2.54
Math and French 2nd grade (pre)	333	0.07	1.03	-2.89	3.00
2012 Math score	318	0.28	0.81	-2.01	2.75
2012 French score	312	0.28	0.88	-1.76	2.13
2012 Math and French score	310	0.31	0.83	-2.37	3.03
Height in 2012	333	160.17	7.91	142.00	180.00
Female	333	0.54	0.50	0.00	1.00
Age 2012	333	21.85	1.39	19.00	26.00
Mother's education	333	5.62	3.65	0.00	17.00
Father's education	333	6.21	3.92	0.00	17.00
Assets 2nd grade	333	-0.08	0.79	-0.76	3.26

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All test scores are constructed using country–specific IRT. Mothers and fathers education were not measured at baseline, but in 2012; hence, they are not reported in this table. Household asset index is constructed using factor analysis. Teacher education index is constructed using factor analysis. The variable consists of variables denoting the education level of the teacher, whether they have formal teaching training, and whether they have done any internships. Teachers and directors experience variables denote the years of experience they have been a teacher and a school director, respectively.

Table A2.2: Mean comparison across panel and full sample of students at baseline – Senegal

SENEGAL	N(Not Panel)	Mean(Not panel)	N(Panel)	Mean(panel)	Difference
French pre 2nd irt 2pl	1424	0.02	448	-0.07	0.09*
French post 2nd irt 2pl	1424	0.02	448	-0.08	0.11**
Math pre 2nd irt 2pl	1424	0.02	448	-0.08	0.11**
Math post 2nd irt 2pl	1424	0.01	448	-0.07	0.08
French Math post 2nd irt 2pl	1424	0.02	448	-0.08	0.10**
French Math pre 2nd irt 2pl	1424	0.02	448	-0.05	0.07
Teacher's education	1375	12.85	420	12.80	0.04
School Infrastructure 2nd grade	1414	0.09	421	-0.15	0.25***
Assets 2nd grade	1427	0.09	448	-0.29	0.38***
Female 1995-96	1425	0.46	448	0.41	0.05*
Age 2nd grade	1407	8.28	444	8.33	-0.05
Teacher education (PCA)	1335	-0.06	400	0.08	-0.14*
Teacher Experience (yrs)	1420	12.70	435	13.48	-0.78
Dir– Years Exp	1397	11.59	418	11.08	0.51

MADAGASCAR	N(Not Panel)	Mean(Not panel)	N(Panel)	Mean(panel)	Difference
French 2nd grade (pre)	-0.01	2044	0.10	333	-0.11*
French 2nd grade (post)	0.02	2044	-0.09	333	0.11*
Math 2nd grade (pre)	-0.01	2044	0.06	333	-0.06
Math 2nd grade (post)	0.01	2044	0.01	333	-0.00
Math and French 2nd grade (pre)	-0.01	2044	0.07	333	-0.08
Math and French 2nd grade (post)	0.01	2044	-0.04	333	0.06
Female	0.51	2005	0.53	331	-0.03
Age second grade	8.74	1919	8.21	324	0.53***
Assets 2nd grade	0.02	2044	-0.08	333	0.10**
Teacher education index	-0.04	1881	0.02	316	-0.06
Teacher experience (yrs)	14.44	1970	13.65	327	0.79
Director experience (yrs)	12.30	1954	13.12	323	-0.82

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. All test scores are constructed using country–specific IRT. Mothers and fathers education were not measured at baseline, but in 2012; hence, they are not reported in this table. Household asset index is constructed using factor analysis. Teacher education index is constructed using factor analysis. The variable consists of variables denoting the education level of the teacher, whether they have formal teaching training, and whether they have done any internships. Teachers and directors experience variables denote the years of experience they have been a teacher and a school director, respectively.

Table A2.3: Mean comparison across clusters chosen for follow–up and full sample of students at baseline

SENEGAL	N(Not followed)	Mean(Not Followed)	N(In follow-up)	Mean(In follow-up)	Difference
French pre 2nd irt 2pl	738	0.07	1134	-0.04	0.11***
French post 2nd irt 2pl	738	0.15	1134	-0.10	0.25***
Math pre 2nd irt 2pl	738	0.12	1134	-0.08	0.21***
Math post 2nd irt 2pl	738	0.14	1134	-0.10	0.23***
French Math post 2nd irt 2pl	738	0.16	1134	-0.11	0.27***
French Math pre 2nd irt 2pl	738	0.12	1134	-0.08	0.20***
Teacher's education	718	12.72	1077	12.92	-0.20*
School Infrastructure 2nd grade	738	0.28	1097	-0.13	0.41***
Assets 2nd grade	738	0.45	1137	-0.29	0.74***
Female 1995-96	738	0.49	1135	0.42	0.07***
Age 2nd grade	720	8.29	1131	8.30	-0.01
Teacher education (PCA)	698	-0.25	1037	0.12	-0.38***
Teacher Experience (yrs)	738	14.22	1117	12.01	2.21***
Dir– Years Exp	738	12.96	1077	10.45	2.52***
MADAGASCAR	N(Not followed)	Mean(Not Followed)	N(In follow-up)	Mean(In follow-up)	Difference
French 2nd grade (pre)	-0.02	1437	0.03	940	-0.04
French 2nd grade (post)	0.07	1437	-0.10	940	0.17***
Math 2nd grade (pre)	-0.03	1437	0.06	940	-0.09**
Math 2nd grade (post)	0.02	1437	-0.02	940	0.04
Math and French 2nd grade (pre)	-0.02	1437	0.04	940	-0.06
Math and French 2nd grade (post)	0.06	1437	-0.08	940	0.14***
Female	0.49	1409	0.55	927	-0.06***
Age second grade	8.61	1349	8.75	894	-0.14*
Assets 2nd grade	0.03	1437	-0.03	940	0.06
Teacher education index	-0.04	1317	-0.02	880	-0.02
Teacher experience (yrs)	14.60	1377	13.91	920	0.68*
Director experience (yrs)	12.29	1377	12.60	900	-0.31

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. The sample consists of all the individuals in the 199 communities in Senegal and 47 in Madagascar that were chosen for the follow–up study, splitting between those in the panel and not in the panel. All test scores are constructed using country–specific IRT. Mothers and fathers education were not measured at baseline, but in 2012; hence, they are not reported in this table. Household asset index is constructed using factor analysis. Teacher education index is constructed using factor analysis. The variable consists of variables denoting the education level of the teacher, whether they have formal teaching training, and whether they have done any internships. Teachers and directors experience variables denote the years of experience they have been a teacher or a school director, respectively.

Table A2.4: Mean comparison within clusters that were chosen for follow–up: individuals reached (panel) and not reached (not in panel) at baseline – Senegal

SENEGAL	N(Not in 2012)	Mean(Not in 2012)	N(In panel)	Mean(In panel)	Difference
French pre 2nd irt 2pl	686	-0.03	448	-0.07	0.04
French post 2nd irt 2pl	686	-0.11	448	-0.08	-0.03
Math pre 2nd irt 2pl	686	-0.08	448	-0.08	-0.00
Math post 2nd irt 2pl	686	-0.12	448	-0.07	-0.05
French Math post 2nd irt 2pl	686	-0.13	448	-0.08	-0.05
French Math pre 2nd irt 2pl	686	-0.09	448	-0.05	-0.04
Teacher's education	657	12.99	420	12.80	0.18
School Infrastructure 2nd grade	676	-0.11	421	-0.15	0.04
Assets 2nd grade	689	-0.29	448	-0.29	-0.00
Female 1995–96	687	0.43	448	0.41	0.02
Age 2nd grade	687	8.28	444	8.33	-0.06
Teacher education (PCA)	637	0.15	400	0.08	0.08
Teacher Experience (yrs)	682	11.07	435	13.48	-2.41***
Dir– Years Exp	659	10.05	418	11.08	-1.02**

MADAGASCAR	N(Not in 2012)	Mean(Not in 2012)	N(In panel)	Mean(In panel)	Difference
French 2nd grade (pre)	-0.01	607	0.10	333	-0.10
French 2nd grade (post)	-0.10	607	-0.09	333	-0.01
Math 2nd grade (pre)	0.06	607	0.06	333	-0.00
Math 2nd grade (post)	-0.03	607	0.01	333	-0.04
Math and French 2nd grade (pre)	0.02	607	0.07	333	-0.05
Math and French 2nd grade (post)	-0.09	607	-0.04	333	-0.05
Female	0.55	596	0.53	331	0.02
Age second grade	9.05	570	8.21	324	0.84***
Assets 2nd grade	-0.00	607	-0.08	333	0.08
Teacher education index	-0.05	564	0.02	316	-0.07
Teacher experience (yrs)	14.06	593	13.65	327	0.41
Director experience (yrs)	12.31	577	13.12	323	-0.81

Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. The sample consists of all the individuals in the 199 communities in Senegal and 47 in Madagascar that were chosen for the follow–up study, splitting between those in the panel and not in the panel. All test scores are constructed using country–specific IRT. Mothers and fathers education were not measured at baseline, but in 2012; hence, they are not reported in this table. Household asset index is constructed using factor analysis. Teacher education index is constructed using factor analysis. The variable consists of variables denoting the education level of the teacher, whether they have formal teaching training, and whether they have done any internships. Teachers and directors experience variables denote the years of experience they have been a teacher or a school director, respectively.

Table A2.5: Grade completed and test scores in 2012 as a function of second grade composite French and math scores: Inverse Probability Weights using the full PASEC baseline sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SENEGAL	Edu Years	Edu Years	Edu Years	Edu Years	Edu Years	Composite	Math	French
	No School FE	School FE	School FE	School FE	School FE	School FE	School FE	School FE
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Second Grade Composite Score	1.801***	2.006***	1.927***	1.886***	1.639***	0.284***	0.599***	0.227***
	(0.194)	(0.213)	(0.196)	(0.197)	(0.328)	(0.070)	(0.118)	(0.069)
Total Obs.	447	447	447	447	447	381	447	381
R–squared	0.168	0.377	0.430	0.435	0.259	0.189	0.218	0.165
F–stat (instrument)					204.7	194.7	204.7	194.7

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
MADAGASCAR	Edu Years	Edu Years	Edu Years	Edu Years	Edu Years	Composite	Math	French
	No School FE	School FE	School FE	School FE	School FE	School FE	School FE	School FE
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Second Grade Composite Score	0.951***	0.385	0.417	0.422	1.311**	0.352**	0.416***	0.243
	(0.244)	(0.318)	(0.277)	(0.277)	(0.541)	(0.148)	(0.161)	(0.153)
Total Obs.	333	333	333	333	333	310	318	312
R–squared	0.083	0.411	0.514	0.516	0.144	0.048	0.006	0.095
F–stat (instrument)					65.8	43.28	44.24	42.84

All regressions weighted with inverse probability weights using the full PASEC sample. Second grade denotes 1995–96 in the case of Senegal and 1997–98 in Madagascar. The sample sizes for the inverse probability weights are 1875 in Senegal, and 2,377 in Madagascar. All test scores are constructed using country–specific IRT. Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

Table A2.6: Grade completed and test scores in 2012 as a function of second–grade composite French and math scores: Inverse Probability Weights using the subset of communities chosen for follow–up

SENEGAL	(1) Edu Years	(2) Edu Years	(3) Edu Years	(4) Edu Years	(5) Edu Years	(6) Composite	(7) Math	(8) French
	No School FE OLS	School FE OLS	School FE OLS	School FE OLS	School FE IV	School FE IV	School FE IV	School FE IV
Second Grade Composite Score	1.671***	1.842***	1.788***	1.754***	1.465***	0.279***	0.568***	0.223***
	(0.188)	(0.206)	(0.195)	(0.197)	(0.309)	(0.068)	(0.116)	(0.066)
Total Obs.	447	447	447	447	447	381	447	381
R–squared	0.151	0.363	0.420	0.425	0.239	0.172	0.199	0.148
F–stat (instrument)					245.4	232.1	245.4	232.1

MADAGASCAR	(1) Edu Years	(2) Edu Years	(3) Edu Years	(4) Edu Years	(5) Edu Years	(6) Composite	(7) Math	(8) French
	No School FE OLS	School FE OLS	School FE OLS	School FE OLS	School FE IV	School FE IV	School FE IV	School FE IV
Second Grade Composite Score	0.982***	0.618*	0.589**	0.595**	1.483***	0.374**	0.446***	0.281***
	(0.226)	(0.318)	(0.267)	(0.267)	(0.501)	(0.153)	(0.166)	(0.155)
Total Obs.	333	333	333	333	333	310	318	312
R–squared	0.082	0.387	0.505	0.508	0.175	0.077	0.024	0.111
F–stat (instrument)					81.98	51.55	53.66	51.04

Height is reported in centimeters. Age is reported in years. Mothers and fathers education are continuous variables measured in years for Madagascar, and dummies for Senegal for any education. Household asset index is constructed using factor analysis. The row widstat denotes the Kleibergen–Paap Wald rk F statistic for weak instruments. Heteroscedasticity–robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.10.

APPENDIX C

APPENDIX – CHAPTER 3

Table A3.1: Mean Differences: Woman–level Analysis Sample

Outcome	In woman–level analysis	Full Sample	Difference
Marriage Age (in years)	17.38	17.40	-0.02
Age (in years)	29.66	29.77	-0.11
Age at first sex (in years)	16.24	16.29	-0.04
Multiple wives dummy	0.25	0.26	-0.01
Wealth Index (0–1)	-0.00	-0.01	0.00
HH Size	5.64	5.63	0.01
Urban Dummy	0.28	0.28	-0.01
Altitude	1310.4	1309.04	1.35
East	0.26	0.26	0.00
North	0.15	0.15	0.00
West	0.28	0.27	0.01
SAMPLE SIZE	4949	5643	

The total sample consists of 5643 women for whom a marriage age is reported. This table compares the mean outcomes for women in the woman–level analysis with the women in the full sample.

Table A3.2: Mean Differences: Woman–level Analysis Sample

Outcome	In woman analysis	Not in woman analysis	Difference
Marriage Age (in years)	17.38	17.55	-0.17
Age (in years)	29.66	29.57	0.09
Age at first sex (in years)	16.24	16.30	-0.06
Multiple wives dummy	0.25	0.31	-0.06*
Wealth Index (0–1)	-0.00	-0.01	0.01
HH Size	5.64	5.59	0.05
Urban Dummy	0.28	0.30	-0.02
Altitude	1310.4	1299.38	11.02
East	0.26	0.25	0.01
North	0.15	0.17	0.02
West	0.28	0.21	0.07**
SAMPLE SIZE	4949	694	

The total sample consists of 5643 women for whom a marriage age is reported. This table compares the mean outcomes for women in the woman–level analysis with the women who are not in the woman–level analysis.

Table A3.3: Mean Differences – Child Analysis Sample

Outcome	In child-level analysis	Full Sample	Difference
Marriage Age (in years)	17.28	17.40	-0.12
Age (in years)	29.56	29.77	-0.21
Age at first sex (in years)	16.22	16.29	-0.07
Multiple wives dummy	0.24	0.26	-0.02
Wealth Index (0–1)	-0.00	-0.01	-0.00
HH Size	5.62	5.63	-0.01
Urban Dummy	0.27	0.28	-0.01
Altitude	1309.4	1309.04	0.35
East	0.26	0.26	0.00
North	0.15	0.15	-0.00
West	0.29	0.27	0.02
SAMPLE SIZE	3998	5643	

The total sample consists of 5643 women for whom a marriage age is reported. This table compares the mean outcomes for women whose children are in the child–level analysis with the women in the full sample.

Table A3.3: Mean Differences – Child Analysis Sample

Outcome	In child analysis	Not in child analysis	Difference
Marriage Age (in years)	17.28	17.52	-0.24
Age (in years)	29.56	29.62	-0.06
Age at first sex (in years)	16.22	16.32	-0.10
Multiple wives dummy	0.24	0.31	-0.07*
Wealth Index (0–1)	-0.00	-0.01	0.01
HH Size	5.62	5.61	0.09
Urban Dummy	0.27	0.31	-0.04
Altitude	1309.4	1301.38	8.02
East	0.26	0.25	0.01
North	0.15	0.17	-0.02
West	0.29	0.24	0.05**
SAMPLE SIZE	3998	1645	

The total sample consists of 5643 women for whom a marriage age is reported.This table compares the mean outcomes for women whose children are in the child–level analysis with the women whose children are not in the child–level analysis.

Table A3.4: Effect of Age at First Sex on Women Outcomes

	Education	Literacy	Employed	Contraception	Age FB	ANC Usage
Age at First Sex	0.18	0.019	0.0022	0.015	0.08	0.008
	(0.10)	(0.012)	(0.010)	(0.012)	(0.068)	(0.075)
Marriage Age	0.32***	0.05***	0.02***	0.02**	0.79***	0.07*
	(0.03)	(0.01)	(0.05)	(0.01)	(0.01)	(0.07)
Total Obs.	4912	4899	4920	4925	4916	4522

The coefficients are from the second stage of a 2SLS IV estimation where age at first sex is instrumented using age at menarche. Marriage age is included as a covariate in the model along with the following controls – woman's height, cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area. The specifications also consist of district and year of birth dummy variables. The standard errors are robust and clustered at the district level. Following are the definitions of the outcome variables: Education (highest grade attained), Literacy (=1 if literate), Labor (=1 if part of the labour force), Contraception (=1 if used contraceptive), Age FB (woman's age at first birth), ANC Usage (=1 if reported using ante natal care).

Table A3.5: Effect of Age at First Sex on Child Outcomes

	HFA z-score	WFA Z-score	BMI Z-score	Hemoglobin	Anemic	Severely Anemic
Age at First Sex	1.11	0.12	1.55	1.10	-0.15	-0.14
	(1.31)	(0.51)	(1.48)	(1.61)	(0.21)	(0.14)
Marriage Age	0.03	0.08	0.06	0.48*	-0.06**	-0.06
	(0.05)	(0.06)	(0.06)	(0.28)	(0.03)	(0.04)
Total Obs.	4997	4982	4990	4956	4958	4958

The coefficients are from the second stage of a 2SLS IV estimation where age at first sex is instrumented using age at menarche. Marriage age is included as a covariate in the model along with the following controls - age of the child, mother's age at the child's birth, child gender and birth order of the child, woman's age at first sex (intercourse), cluster altitude (in meters), household size, wealth index and categorical variables for the presence of multiple wives, religion, ethnicity, and living in an urban area. The specifications also consist of district and year of birth (dummy variables. The standard errors are robust and clustered at the district level. Following are the definitions of the outcome variables: HFA (Height for Age Z-score), WFA (Weight for Age Z-Score), BMI Z-score, Hemoglobin levels (g/dl), Anemic (=1 if hemoglobin below 11 g/dl) and Severely Anemic (=1 if hemoglobin below 7 g/dl). The coefficients are from the second stage of a 2SLS IV estimation. The standard errors are robust and clustered at the district level.

BIBLIOGRAPHY

Ran Abramitzky, Adeline Delavande, and Luis Vasconcelos. Marrying up: the role of sex ratio in assortative matching. *American Economic Journal: Applied Economics*, 3(3):124–157, 2011.

Linda S Adair. Size at birth predicts age at menarche. *Pediatrics*, 107(4):e59–e59, 2001.

Linda S Adair. Size at birth and growth trajectories to young adulthood. *American Journal of Human Biology*, 19(3):327–337, 2007.

IO Adebara and Munir'deen Ijaiya. Relationship of biosocial factors and hemoglobin genotype with age of menarche among school girls in ilorin, nigeria. *Nepal Journal of Obstetrics and Gynaecology*, 7(1), 2013.

Anjali Adukia. Sanitation and education. *American Economic Journal: Applied Economics*, 9(2):23–59, 2017.

Jorge M Agüero and Prashant Bharadwaj. Do the more educated know more about health? evidence from schooling and hiv knowledge in zimbabwe. *Economic Development and Cultural Change*, 62(3):489–517, 2014.

Dennis Ahlburg. Intergenerational transmission of health. *The American Economic Review*, 88(2):265–270, 1998.

Harold Alderman and Derek D Headey. How important is parental education for child nutrition? *World Development*, 94:448–464, 2017.

Harold Alderman, Jere R Behrman, Hans-Peter Kohler, John A Maluccio, and Susan Cotts Watkins. Attrition in longitudinal household survey data: Some tests for three developing-country samples. *Demographic Research*, 5:79–124, 2001.

Harold Alderman, Jooseop Kim, and Peter F Orazem. Design, evaluation, and sustainability of private schools for the poor: the pakistan urban and rural fellowship school experiments. *Economics of Education Review*, 22(3):265–274, 2003.

Harold Alderman, John Hoddinott, and Bill Kinsey. Long term consequences of early childhood malnutrition. *Oxford Economic Papers*, 58(3):450–474, 2006.

Douglas Almond. Is the 1918 influenza pandemic over? long-term effects of in utero influenza exposure in the post-1940 us population. *Journal of Political Economy*, 114(4):672–712, 2006.

Douglas Almond and Janet Currie. Killing me softly: The fetal origins hypothesis. *The Journal of Economic Perspectives*, 25(3):153–172, 2011a.

Douglas Almond and Janet Currie. Killing me softly: The fetal origins hypothesis. *The Journal of Economic Perspectives*, 25(3):153, 2011b.

Joseph G Altonji, Todd E Elder, and Christopher R Taber. An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *Journal of Human Resources*, 40(4):791–821, 2005.

Vikesh Amin, Jere R Behrman, and Tim D Spector. Does more schooling improve health outcomes and health related behaviors? evidence from uk twins. *Economics of Education Review*, 35:134–148, 2013.

Mabel Andalón, Jenny Williams, and Michael Grossman. Empowering women: The effect of schooling on young women's knowledge and use of contraception. Technical report, National Bureau of Economic Research, 2014.

Siwan Anderson. The economics of dowry and brideprice. *Journal of Economic Perspectives*, 21(4):151–174, 2007.

Tahir Andrabi, Jishnu Das, and Asim Ijaz Khwaja. What did you do all day? maternal education and child outcomes. *Journal of Human Resources*, 47(4): 873–912, 2012.

Tahir Andrabi, Jishnu Das, and Asim Ijaz Khwaja. *Report cards: The impact of providing school and child test scores on educational markets*. The World Bank, 2015.

Joshua Angrist, Eric Bettinger, and Michael Kremer. Long-term educational consequences of secondary school vouchers: Evidence from administrative records in colombia. *American Economic Review*, 96(3):847–862, 2006.

Joshua D. Angrist and Alan B. Krueger. Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments. *Journal of Economic Perspectives*, 15(4):69–85, 2001.

Joshua D Angrist, Parag A Pathak, and Christopher R Walters. Explaining charter school effectiveness. *American Economic Journal: Applied Economics*, 5(4): 1–27, 2013.

ASER. Enrolment and learning report card. 2016.

Nava Ashraf, Dean Karlan, and Wesley Yin. Female empowerment: Impact of a commitment savings product in the philippines. *World Development*, 38(3): 333–344, 2010.

M Aslam and G Kingdon. Parental education and child health: understanding the pathways of impact in pakistan. centre for the study of african economies working paper series (csaewps) 2010-16, 2010.

F Aubery and D Sahn. Cognitive achievement production in madagascar: A value-added model approach. *Cornell University, Centre d'Etude et de Recherche sur le Développement International (CERDI)*, 2014.

Mehtabul Azam and Chan Hang Saing. Assessing the impact of district primary education program in india. *Review of Development Economics*, 21(4):1113–1131, 2017.

Theophile T Azomahou and Eleni A Yitbarek. *Intergenerational education mobility in Africa: has progress been inclusive?* The World Bank, 2016.

Sarah Baird, Craig McIntosh, and Berk Özler. Cash or condition? evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, 126(4):1709–1753, 2011.

Oriana Bandiera, Robin Burgess, Markus Goldstein, Niklas Buehren, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. Womens' empowerment in action: evidence from a randomized control trial in africa. 2014.

Abhijit V Banerjee, Shawn Cole, Esther Duflo, and Leigh Linden. Remedying education: Evidence from two randomized experiments in india. *The Quarterly Journal of Economics*, 122(3):1235–1264, 2007.

Rukmini Banerji, James Berry, and Marc Shotland. The impact of maternal literacy and participation programs: Evidence from a randomized evaluation in india. *American Economic Journal: Applied Economics*, 9(4):303–37, 2017.

Sarah L Barber and Paul J Gertler. Empowering women: how mexico's conditional cash transfer programme raised prenatal care quality and birth weight. *Journal of Development Effectiveness*, 2(1):51–73, 2010.

David J Barker. Fetal origins of coronary heart disease. *BMJ: British Medical Journal*, 311(6998):171, 1995.

Felipe Barrera-Osorio, Marianne Bertrand, Leigh L Linden, and Francisco Perez-Calle. Improving the design of conditional transfer programs: Evidence from a randomized education experiment in colombia. *American Economic Journal: Applied Economics*, 3(2):167–95, 2011.

Robert J Barro and Jong Wha Lee. A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104:184–198, 2013.

Alaka Malwade Basu and Gayatri Brij Koolwal. Two concepts of female empowerment: Some leads from dhs data on womens status and reproductive health. *A focus on gender–Collected papers on gender using DHS data, ORC Macro, Calverton MD*, pages 15–33, 2005.

Gary S Becker. A theory of marriage: Part i. *Journal of Political Economy*, 81(4): 813–846, 1973.

Gary S Becker. A treatise on the family, enl. ed. *Cambridge, Mass: Harvard*, 1991.

Gary S Becker and H Gregg Lewis. On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2, Part 2):S279–S288, 1973.

Kathleen Beegle, Elizabeth Frankenberg, and Duncan Thomas. Bargaining power within couples and use of prenatal and delivery care in indonesia. *Studies in Family Planning*, 32(2):130–146, 2001.

Jere R Behrman and Mark R Rosenzweig. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review*, 92(1): 323–334, 2002.

Jere R Behrman and Mark R Rosenzweig. Returns to birthweight. *Review of Economics and Statistics*, 86(2):586–601, 2004.

Jere R Behrman, David Ross, and Richard Sabot. Improving quality versus increasing the quantity of schooling: Estimates of rates of return from rural pakistan. *Journal of Development Economics*, 85(1-2):94–104, 2008.

Jere R Behrman, Susan W Parker, and Petra E Todd. Schooling impacts of conditional cash transfers on young children: Evidence from mexico. *Economic Development and Cultural Change*, 57(3):439–477, 2009.

Jere R Behrman, John Hoddinott, John A Maluccio, Erica Soler-Hampejsek, Emily L Behrman, Reynaldo Martorell, Manuel Ramírez-Zea, and Aryeh D Stein. What determines adult cognitive skills? influences of pre-school, school, and post-school experiences in guatemala. *Latin American Economic Review*, 23(1):4, 2014.

Jere R Behrman, Whitney Schott, Subha Mani, Benjamin T Crookston, Kirk Dearden, Le Thuc Duc, Lia CH Fernald, and Aryeh D Stein. Intergenerational transmission of poverty and inequality: parental resources and schooling attainment and childrens human capital in ethiopia, india, peru, and vietnam. *Economic Development and Cultural Change*, 65(4):657–697, 2017.

Julia Andrea Behrman. Does schooling affect womens desired fertility? evidence from malawi, uganda, and ethiopia. *Demography*, 52(3):787–809, 2015.

Peter Bergman. Parent-child information frictions and human capital investment: Evidence from a field experiment. 2015.

Catherine S. Berkey, Jane D. Gardner, A. Lindsay Frazier, and Graham A. Colditz. Relation of childhood diet and body size to menarche and adolescent growth in girls. *American Journal of Epidemiology*, 152(5):446–452, 2000.

Eric Bettinger and Robert Slonim. Patience among children. *Journal of Public Economics*, 91(1-2):343–363, 2007.

Sonia Bhalotra and Samantha Rawlings. Gradients of the intergenerational transmission of health in developing countries. *Review of Economics and Statistics*, 95(02):660–672, 2013.

Prashant Bharadwaj, Juan Eberhard, and Christopher Neilson. Do initial endowments matter only initially? the persistent effect of birth weight on school achievement. 2010.

Paul Bingley, Kaare Christensen, and Ian Walker. Twin-based estimates of the returns to education: evidence from the population of danish twins. *Working Paper*, 2005.

Paul Bingley, Kaare Christensen, and Vibeke Myrup Jensen. Parental schooling and child development: Learning from twin parents. 2009.

ALord Birnbaum. Some latent trait models and their use in inferring an examinee's ability. *Statistical Theories of Mental Test Scores*, 1968.

David Bishai and Shoshana Grossbard. Far above rubies: Bride price and extramarital sexual relations in uganda. *Journal of Population Economics*, 23(4): 1177–1187, 2010. ISSN 0933-1433. doi: 10.1007/s00148-008-0226-3.

Martina Björkman-Nyqvist. Income shocks and gender gaps in education: Evidence from uganda. *Journal of Development Economics*, 105:237–253, 2013.

Martina Björkman Nyqvist and Seema Jayachandran. Mothers care more, but fathers decide: Educating parents about child health in uganda. *American Economic Review*, 107(5):496–500, 2017.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review*, 95(1):437–449, 2005.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1):409–439, 2007.

Moussa P Blimpo. Team incentives for education in developing countries: A randomized field experiment in benin. *American Economic Journal: Applied Economics*, 6(4):90–109, 2014.

Francis Bloch, Vijayendra Rao, and Sonalde Desai. Wedding celebrations as conspicuous consumption signaling social status in rural india. *Journal of Human Resources*, 39(3):675–695, 2004.

Emilia Del Bono, Marco Francesconi, Yvonne Kelly, and Amanda Sacker. Early maternal time investment and early child outcomes. *The Economic Journal*, 126 (596):F96–F135, 2016.

Thomas Bossuroy and Denis Cogneau. Social mobility in five a frican countries. *Review of Income and Wealth*, 59:S84–S110, 2013.

John Bound, David A. Jaeger, and Regina M. Baker. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430):443–450, 1995.

Victoria J Bourne, Helen C Fox, Ian J Deary, and Lawrence J Whalley. Does childhood intelligence predict variation in cognitive change in later life? *Personality and Individual Differences*, 42(8):1551–1559, 2007.

Nina Buchmann, Erica Field, Rachel Glennerster, Shahana Nazneen, Svetlana Pimkina, and Iman Sen. Power vs money: Alternative approaches to reducing child marriage in bangladesh, a randomized control trial. 2017.

Niklas Buehren, Markus Goldstein, Selim Gulesci, Munshi Sulaiman, and Venus Yam. Evaluation of layering microfinance on an adolescent development program for girls in tanzania. *Working Paper*, 2015.

Dana Burde and Leigh L Linden. Bringing education to afghan girls: A randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics*, 5(3):27–40, 2013.

Lizette Burgers. Background and rationale for school sanitation and hygiene education. *New York: UNICEF*, 2000.

Anne Burrows 1 and Sally Johnson. Girls' experiences of menarche and menstruation. *Journal of Reproductive and Infant Psychology*, 23(3):235–249, 2005.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82 (6):2295–2326, 2014.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. Coverage error optimal confidence intervals for regression discontinuity designs. *Working Paper*, 2016.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association*, (just-accepted), 2017.

Pedro Carneiro, Costas Meghir, and Matthias Parey. Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, 11(suppl_1):123–160, 2013.

Anne Case and Christina Paxson. Height, health, and cognitive function at older ages. *The American Economic Review*, 98(2):463, 2008a.

Anne Case and Christina Paxson. Stature and status: Height, ability, and labor market outcomes. *Journal of Political Economy*, 116(3):499–532, 2008b.

Anne Case and Christina Paxson. Causes and consequences of early-life health. *Demography*, 47(1):S65–S85, 2010.

Anne Case, Darren Lubotsky, and Christina Paxson. Economic status and health in childhood: The origins of the gradient. *American Economic Review*, 92(5): 1308–1334, 2002.

Anne Case, Angela Fertig, and Christina Paxson. The lasting impact of childhood health and circumstance. *Journal of Health Economics*, 24(2):365–389, 2005.

Matias D Cattaneo and Gonzalo Vazquez-Bare. The choice of neighborhood in regression discontinuity designs. *Observational Studies*, 2:134Ā146, 2016.

AV Chari, Rachel Heath, Annemie Maertens, and Freeha Fatima. The causal effect of maternal age at marriage on child wellbeing: Evidence from india. *Journal of Development Economics*, 127:42–55, 2017.

Yuk F Cheong, Kathryn M Yount, and Alice Ann Crandall. Longitudinal measurement invariance of the womens agency scale. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 134(1):24–36, 2017.

Arnaud Chevalier and Tarja K Viitanen. The long-run labour market consequences of teenage motherhood in britain. *Journal of Population Economics*, 16 (2):323–343, 2003.

Arnaud Chevalier, Colm Harmon, Vincent OSullivan, and Ian Walker. The impact of parental income and education on the schooling of their children. *IZA Journal of Labor Economics*, 2(1):8, 2013.

Shin-Yi Chou, Jin-Tan Liu, Michael Grossman, and Ted Joyce. Parental education and child health: evidence from a natural experiment in taiwan. *American Economic Journal: Applied Economics*, 2(1):33–61, 2010.

S Chowdhury, AK Shahabuddin, AJ seal, KK Talukder, Q Hassan, RA Begum, Q Rahman, A Tomkins, A Costello, and MQ Talukder. Nutritional status and age at menarche in a rural area in bangladesh. *Annals of Human Biology*, 27(3): 249–256, 2000.

Luc Christiaensen and Harold Alderman. Child malnutrition in ethiopia: can maternal knowledge augment the role of income? *Economic Development and Cultural Change*, 52(2):287–312, 2004.

Shelley Clark, Judith Bruce, and Annie Dude. Protecting young women from hiv/aids: the case against child and adolescent marriage. *International family planning perspectives*, pages 79–88, 2006.

Katja Coneus and C Katharina Spiess. The intergenerational transmission of health in early childhoodevidence from the german socio-economic panel study. *Economics & Human Biology*, 10(1):89–97, 2012.

Timothy G Conley, Christian B Hansen, and Peter E Rossi. Plausibly exogenous. *Review of Economics and Statistics*, 94(1):260–272, 2012.

AliceAnn Crandall, Kristin VanderEnde, Yuk Fai Cheong, Sylvie Dodell, and Kathryn M Yount. Women's age at first marriage and postmarital agency in egypt. *Social Science Research*, 57:148–160, 2016.

Flavio Cunha and James Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, 2007.

Flavio Cunha and James J Heckman. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4):738–782, 2008.

Flavio Cunha, James J Heckman, and Susanne M Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3): 883–931, 2010.

Anne E Cunningham and Keith E Stanovich. Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental psychology*, 33(6):934, 1997.

Janet Currie and Rosemary Hyson. Is the impact of health shocks cushioned by socioeconomic status? the case of low birthweight. *American Economic Review*, 89(2):245–250, 1999.

Janet Currie and Brigitte C Madrian. Health, health insurance and the labor market. *Handbook of Labor Economics*, 3:3309–3416, 1999.

Janet Currie and Enrico Moretti. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532, 2003.

Janet Currie and Tom Vogl. Early-life health and adult circumstance in developing countries. *Annual Review of Economics*, 5(1):1–36, 2013.

Gordon B Dahl. Early teen marriage and future poverty. *Demography*, 47(3): 689–718, 2010.

Jishnu Das and Tristan Zajonc. India shining and bharat drowning: Comparing two indian states to the worldwide distribution in mathematics achievement. *Journal of Development Economics*, 2(92):175–187, 2010.

Jacobus De Hoop and Furio C Rosati. Does promoting school attendance reduce child labor? evidence from burkina faso's bright project. *Economics of Education Review*, 39:78–96, 2014.

Marleen Dekker and Hans Hoogeveen. Bride wealth and household security in rural zimbabwe. *Journal of African Economies*, 11(1):114–145, 2002.

David J Deming. Using school choice lotteries to test measures of school effectiveness. *American Economic Review*, 104(5):406–11, 2014.

David J Deming, Justine S Hastings, Thomas J Kane, and Douglas O Staiger. School choice, school quality, and postsecondary attainment. *American Economic Review*, 104(3):991–1013, 2014.

Stefan Dercon and Catherine Porter. Live aid revisited: Long-term impacts of the 1984 ethiopian famine on children. *Journal of the European Economic Association*, 12(4):927–948, 2014.

Sonalde Desai and Soumya Alva. Maternal education and child health: Is there a strong causal relationship? *Demography*, 35(1):71–81, 1998.

Sonalde Desai and Reeve Vanneman. National council of applied economic research, new delhi. 2005. *India human development survey (IHDS)*, 2005.

Sonalde Desai, Amaresh Dubey, Reeve Vanneman, Rukmini Banerji, et al. Private schooling in india: A new educational landscape. In *India Policy Forum*, volume 5, pages 1–38. National Council of Applied Economic Research, 2009.

Cheryl Doss. Intrahousehold bargaining and resource allocation in developing countries. *The World Bank Research Observer*, 28(1):52–78, 2013.

Esther Duflo. Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813, 2001.

Esther Duflo. The medium run effects of educational expansion: Evidence from a large school construction program in indonesia. *Journal of Development Economics*, 74(1):163–197, 2004.

Esther Duflo. Women empowerment and economic development. *Journal of Economic Literature*, 50(4):1051–79, 2012.

Esther Duflo, Rema Hanna, and Stephen P Ryan. Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–78, 2012.

Esther Duflo, Pascaline Dupas, and Michael Kremer. Education, hiv, and early fertility: Experimental evidence from kenya. *American Economic Review*, 105 (9):2757–97, 2015.

Christelle Dumas and Sylvie Lambert. Educational achievement and socio-economic background: causality and mechanisms in senegal. *Journal of African Economies*, 20(1):1–26, 2010.

Greg J Duncan, Chantelle J Dowsett, Amy Claessens, Katherine Magnuson, Aletha C Huston, Pamela Klebanov, Linda S Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, et al. School readiness and later achievement. *Developmental psychology*, 43(6):1428, 2007.

Greg J Duncan, Katherine Magnuson, et al. The nature and impact of early achievement skills, attention skills, and behavior problems. *Whither Opportunity*, pages 47–70, 2011.

Martha G Eide, Nina Øyen, Rolv Skjœrven, Stein Tore Nilsen, Tor Bjerkedal, and Grethe S Tell. Size at birth and gestational age as predictors of adult height and weight. *Epidemiology*, pages 175–181, 2005.

Bruce J Ellis. Timing of pubertal maturation in girls: an integrated life history approach. *Psychological Bulletin*, 130(6):920, 2004.

Patrice L Engle. The role of men in families: Achieving gender equity and suppor ting children. *Gender & Development*, 5(2):31–40, 1997.

Marcel Fafchamps and Agnes R Quisumbing. Marriage, bequest, and assortative matching in rural ethiopia. *Economic Development and Cultural Change*, 53 (2):347–380, 2005.

Leon Feinstein. Inequality in the early cognitive development of british children in the 1970 cohort. *Economica*, 70(277):73–97, 2003.

Alicia Fentiman, Andrew Hall, and Donald Bundy. School enrolment patterns in rural ghana: a comparative study of the impact of location, gender, age and health on children's access to basic schooling. *Comparative Education*, 35 (3):331–349, 1999.

Erica Field and Attila Ambrus. Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of Political Economy*, 116(5):pp. 881–930, 2008a.

Erica Field and Attila Ambrus. Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of Political Economy*, 116(5):881–930, 2008b.

Mario Fiorini and Michael P Keane. How the allocation of childrens time affects cognitive and noncognitive development. *Journal of Labor Economics*, 32(4):787–836, 2014.

Robert W Fogel. The conquest of high mortality and hunger in europe and america: Timing and mechanisms, 1990.

Robert William Fogel. New sources and new techniques for the study of secular trends in nutritional status, health, mortality, and the process of aging. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 26(1):5–43, 1993.

David S Freedman, Laura Kettel Khan, Mary K Serdula, William H Dietz, Sathanur R Srinivasan, and Gerald S Berenson. Racial differences in the tracking of childhood bmi to adulthood. *Obesity Research*, 13(5):928–935, 2005.

Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, (just-accepted), 2017.

Animesh Giri and Vinish Shrestha. The effect of school construction on educational outcomes among females: Evidence from nepal. *Development Journal of the South*, 2(1):1, 2017.

Paul Glewwe. The relevance of standard estimates of rates of return to schooling for education policy: A critical assessment. *Journal of Development economics*, 51(2):267–290, 1996.

Paul Glewwe. Why does mother's schooling raise child health in developing countries? evidence from morocco. *Journal of Human Resources*, pages 124–159, 1999.

Paul Glewwe and Karthik Muralidharan. Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications. In *Handbook of the Economics of Education*, volume 5, pages 653–743. Elsevier, 2016.

Paul Glewwe, Michael Kremer, and Sylvie Moulin. Many children left behind? textbooks and test scores in kenya. *American Economic Journal: Applied Economics*, 1(1):112–35, 2009.

Paul Glewwe, Sofya Krutikova, and Caine Rolleston. Do schools reinforce or reduce learning gaps between advantaged and disadvantaged students? evidence from vietnam and peru. *Economic Development and Cultural Change*, 65 (4):699–739, 2017.

Peter Glick and David E Sahn. Cognitive skills among children in senegal: Disentangling the roles of schooling and family background. *Economics of Education Review*, 28(2):178–188, 2009.

Peter Glick and David E Sahn. Early academic performance, grade repetition, and school attainment in senegal: A panel data analysis. *The World Bank Economic Review*, 24(1):93–120, 2010.

Peter Glick, Jean Claude Randrianarisoa, and David E Sahn. Family background, school characteristics, and children's cognitive achievement in madagascar. *Education Economics*, 19(4):363–396, 2011.

Peter Glick, Christopher Handy, and David E Sahn. Schooling, marriage, and age at first birth in madagascar. *Population Studies*, 69(2):219–236, 2015.

Srinivas Goli, Anu Rammohan, and Deepti Singh. The effect of early marriages and early childbearing on womens nutritional status in india. *Maternal and Child Health Journal*, 19(8):1864–1880, 2015.

Rubén Grajeda, Jere R Behrman, Rafael Flores, John A Maluccio, Reynaldo Martorell, and Aryeh D Stein. The human capital study 2002–04: tracking, data collection, coverage, and attrition. *Food and Nutrition Bulletin*, 26(2_suppl1): S15–S24, 2005.

Monica J Grant. The demographic promise of expanded female education: trends in the age at first birth in malawi. *Population and Development Review*, 41(3):409–438, 2015.

Karen A Grépin and Prashant Bharadwaj. Maternal education and child mortality in zimbabwe. *Journal of Health Economics*, 44:97–117, 2015.

Michael Grossman. The relationship between health and schooling: Whats new? Technical report, National Bureau of Economic Research, 2015.

Pınar Mine Güneş. The role of maternal education in child health: Evidence from a compulsory schooling law. *Economics of Education Review*, 47:1–16, 2015.

Youjin Hahn, Asadul Islam, Kanti Nuzhat, Russell Smyth, and Hee-Seung Yang. Education, marriage, and fertility: Long-term evidence from a female stipend program in bangladesh. *Economic Development and Cultural Change*, 66(2):383–415, 2018.

Sudhanshu Handa. Raising primary school enrolment in developing countries: The relative importance of supply and demand. *Journal of Development Economics*, 69(1):103–128, 2002.

Eric A Hanushek. Economic growth in developing countries: The role of human capital. *Economics of Education Review*, 37:204–212, 2013.

Eric A Hanushek and Ludger Woessmann. The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–68, 2008.

James Heckman and Edward Vytlacil. Instrumental variables methods for the correlated random coefficient model: Estimating the average rate of return to schooling when the return is correlated with schooling. *Journal of Human Resources*, pages 974–987, 1998.

James J Heckman. Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902, 2006.

James J Heckman and Stefano Mosso. The economics of human development and social mobility. *Annual Review of Economics*, 6(1):689–733, 2014.

Christian Helmers and Manasa Patnam. The formation and evolution of childhood skill acquisition: Evidence from india. *Journal of Development Economics*, 95(2):252–266, 2011.

Joan Hamory Hicks and Daniel L Hicks. Lucky late bloomers? the consequences of early marriage for women in western kenya. 2015.

Helena Holmlund, Mikael Lindahl, and Erik Plug. The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature*, 49(3):615–51, 2011.

Seo Yeon Hong and Leopold Remi Sarr. Long-term impacts of the free tuition and female stipend programs on education attainment, age of marriage, and married women's labor market participation of in bangladesh. *Washington, DC: World Bank. Retrieved from http://documents. worldbank. org/curated/en/2012/10/18257128/long-term-impacts-free-tuition-female-stipend-programs-education-attainment-age-marriage-married-womens-labor-market-participation-bangladesh*, 2012.

V Joseph Hotz, Charles H Mullin, and Seth G Sanders. Bounding causal effects using data from a contaminated natural experiment: Analysing the effects of teenage childbearing. *The Review of Economic Studies*, 64(4):575–603, 1997.

Andrew J Houtenville and Karen Smith Conway. Parental effort, school resources, and student achievement. *Journal of Human Resources*, 43(2):437–453, 2008.

Jennifer Hunt. Teen births keep american crime high. Technical report, National Bureau of Economic Research, 2003.

Suliman Ahmad Huzayyin and George T Acsadi. Family and marriage in some african asiatic countries. *Research Monograph Series*, (6), 1976.

Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959, 2012.

Shayesteh Jahanfar, Munn-Sann Lye, and Isthrinayagy S Krishnarajah. Genetic and environmental effects on age at menarche, and its relationship with reproductive health in twins. *Indian Journal of Human Genetics*, 19(2):245, 2013.

Saranga Jain and Kathleen Kurz. *New insights on preventing child marriage: A global analysis of factors and programs*. International Center for Research on Women (ICRW), 2007.

J Jalan and E Glinskaya. Small bang for big bucks. *An Evaluation of a Primary School Intervention in India. Centre for Training and Research in Public Finance And Policy, Calcutta*, 2013.

Robert Jensen. The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125(2):515–548, 2010.

C Chrisler Joan and PhD Carolyn B Zittel. Menarche stories: Reminiscences of college students from lithuania, malaysia, sudan, and the united states. *Health Care for Women International*, 19(4):303–312, 1998.

Child Health Epidemiology Reference Group Kozuki Naoko Lee Anne C. Katz Joanne. Moderate to severe, but not mild, maternal anemia is associated with increased risk of small-for-gestational-age outcomes–. *The Journal of Nutrition*, 142(2):358–362, 2011.

David W Johnston, Grace Lordan, Michael A Shields, and Agne Suziedelyte. Education and health knowledge: Evidence from uk compulsory schooling reform. *Social Science & Medicine*, 127:92–100, 2015.

Nicola Jones, Bekele Tefera, Janey Stephenson, Taveeshi Gupta, and Paola Pereznieto. Early marriage and education: the complex role of social norms in shaping ethiopian adolescent girls' lives. *London: ODI*, 2014.

Shareen Joshi and T Paul Schultz. Family planning as an investment in development: evaluation of a program's consequences in matlab, bangladesh. 2007.

Naila Kabeer. Education is my daughter's future. *Index on Censorship*, 27(3): 154–160, 1998.

Naila Kabeer. Paid work, women's empowerment and gender justice: critical pathways of social change. 2008.

Naila Kabeer. Between affiliation and autonomy: navigating pathways of women's empowerment and gender justice in rural bangladesh. *Development and Change*, 42(2):499–528, 2011.

Heidi Kaila, David Sahn, and Naveen Sunder. Early life determinants of cognitive ability: A comparative study on madagascar and senegal. *Working Paper*, 2019.

Thomas J Kane, Daniel F McCaffrey, Trey Miller, and Douglas O Staiger. Have we identified effective teachers? validating measures of effective teaching using random assignment. In *Research Paper. MET Project. Bill & Melinda Gates Foundation*. Citeseer, 2013.

AK Kapoor and Satwanti Kapoor. The effects of high altitude on age at menarche and menopause. *International Journal of Biometeorology*, 30(1):21–26, 1986.

Dan K Kaye, Florence Mirembe, A Johansson, AM Ekstrom, and GB Kyomuhendo. Implications of bride price on domestic violence and reproductive health in wakiso district, uganda. *African Health Sciences*, 5(4):300–303, 2005.

Harounan Kazianga, Dan Levy, Leigh L Linden, et al. The effects of. *American Economic Journal: Applied Economics*, 5(3):41–62, 2013.

Kate J Kerber, Joseph E de Graft-Johnson, Zulfiqar A Bhutta, Pius Okong, Ann Starrs, and Joy E Lawn. Continuum of care for maternal, newborn, and child health: from slogan to service delivery. *The Lancet*, 370(9595):1358–1369, 2007.

Gaurav Khanna. Large-scale education reform in general equilibrium: Regression discontinuity evidence from india. *Unpublished Paper, University of Michigan*, 2015.

Elizabeth M King, Stephan Klasen, and Maria Porter. *Women and development*. Copenhagen Consensus Center, 2007.

Mervyn King, Enrique Sentana, and Sushil Wadhwani. Volatility and links between national stock markets. *Econometrica*, 62(4):901–33, 1994.

Geeta Gandhi Kingdon. The progress of school education in india. *Oxford Review of Economic Policy*, 23(2):168–195, 2007.

Jackie Kirk and Marni Sommer. Menstruation and body awareness: linking girls health with girls education. *Royal Tropical Institute (KIT), Special on Gender and Health*, pages 1–22, 2006.

Daniel Klepinger, Shelly Lundberg, and Robert Plotnick. How does adolescent fertility affect the human capital and wages of young women? *Journal of Human Resources*, pages 421–448, 1999.

Zovanga Louis Kone, Maggie Y Liu, Aaditya Mattoo, Caglar Ozden, and Siddharth Sharma. Internal borders and migration in india. 2017.

Michael Kremer, Conner Brannen, and Rachel Glennerster. The challenge of education and learning in the developing world. *Science*, 340(6130):297–300, 2013.

Adriana D Kugler and Santosh Kumar. Preference for boys, family size, and educational attainment in india. *Demography*, 54(3):835–859, 2017.

Hemanshu Kumar and Rohini Somanathan. Mapping indian districts across census years, 1971-2001. *Economic and Political Weekly*, pages 69–73, 2009.

Daniel LaFave and Duncan Thomas. Height and cognition at work: Labor market productivity in a low income setting. *Economics & Human Biology*, 25: 52–64, 2017.

Tjalling Leenstra, LT Petersen, SK Kariuki, AJ Oloo, PA Kager, and FO Ter Kuile. Prevalence and severity of malnutrition and age at menarche; cross-sectional studies in adolescent schoolgirls in western kenya. *European Journal of Clinical Nutrition*, 59(1):41–48, 2005.

Steven D Levitt, John A List, Susanne Neckermann, and Sally Sadoff. The impact of short-term incentives on student performance. *Unpublished mimeo, University of Chicago*, 2011.

Arthur Lewbel. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30 (1):67–80, 2012.

Tao Li, Li Han, Linxiu Zhang, and Scott Rozelle. Encouraging classroom peer interactions: Evidence from chinese migrant schools. *Journal of Public Economics*, 111:29–45, 2014.

Prashant Loyalka, Chengfang Liu, Yingquan Song, Hongmei Yi, Xiaoting Huang, Jianguo Wei, Linxiu Zhang, Yaojiang Shi, James Chu, and Scott Rozelle. Can information and counseling help students from poor rural areas go to high school? evidence from china. *Journal of Comparative Economics*, 41(4):1012–1025, 2013.

Y. M. A. Lubaale. Child marriages in uganda after enacting the death penalty for defilement. *Psychology abd Behavioural Sciences*, 2(3):138–147, 2013.

Shelly Lundberg and Robert A Pollak. Separate spheres bargaining and the marriage market. *Journal of Political Economy*, 101(6):988–1010, 1993.

Petter Lundborg. The health returns to schoolingwhat can we learn from twins? *Journal of Population Economics*, 26(2):673–701, 2013.

Petter Lundborg, Paul Nystedt, and Dan-Olof Rooth. Height and earnings: The role of cognitive and noncognitive skills. *Journal of Human Resources*, 49(1): 141–166, 2014.

Simeen Mahmud, Nirali M Shah, and Stan Becker. Measurement of womens empowerment in rural bangladesh. *World Development*, 40(3):610–619, 2012.

Kaveh Majlesi. Labor market opportunities and women's decision making power within households. *Journal of Development Economics*, 119:34–47, 2016.

Chata Male and Quentin Wodon. Basic profile of child marriage in uganda. health, nutrition and population knowledge brief. *World Bank, Washington, DC.*, 2016.

Francesca Marchetta and David E Sahn. The role of education and family background in marriage, childbearing, and labor market participation in senegal. *Economic Development and Cultural Change*, 64(2):369–403, 2016.

Reynaldo Martorell. Enhancing human potential in guatemalan adults through improved nutrition in early childhood. *Nutrition Today*, 28(1):6–13, 1993.

Reynaldo Martorell and Jean-Pierre Habicht. Growth in early childhood in developing countries. 1986a.

Reynaldo Martorell and Jean-Pierre Habicht. Growth in early childhood in developing countries. 1986b.

Kenji Matsuura and Cort Willmott. Terrestrial air temperature and precipitation: 1900-2014 gridded monthly time series, version 4.01. *University of Delaware.*, 2015.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

Katharina Michaelowa. Primary education quality in francophone sub-saharan africa: Determinants of learning achievement and efficiency considerations. *World Development*, 29(10):1699–1716, 2001.

Beatrice Odongkara Mpora, Thereza Piloya, Sylvia Awor, Thomas Ngwiri, Paul Laigong, Edison A Mworozi, and Zeev Hochberg. Age at menarche in relation to nutritional status and critical life events among rural and urban secondary school girls in post-conflict northern uganda. *BMC Women's Health*, 14(1):66, 2014.

Karthik Muralidharan. Priorities for primary education policy in indias 12th five-year plan. In *India Policy Forum*, volume 9, pages 1–61. National Council of Applied Economic Research, 2013.

Karthik Muralidharan and Michael Kremer. Public and private schools in rural india. *Harvard University, Department of Economics, Cambridge, MA*, 2006.

Karthik Muralidharan and Nishith Prakash. Cycling to school: increasing secondary school enrollment for girls in india. *American Economic Journal: Applied Economics*, 9(3):321–50, 2017.

Karthik Muralidharan, Abhijeet Singh, and Alejandro J Ganimian. Disrupting education? experimental evidence on technology-aided instruction in india. Technical report, National Bureau of Economic Research, 2016.

Nawal M Nour. Health consequences of child marriage in africa. *Emerging Infectious Diseases*, 12(11):1644, 2006.

Martin O'Flaherty, Janeen Baxter, Michele Haynes, and Gavin Turrell. Age at first birth and long-term health in life course perspective: contextualizing the health effects of fertility timing. Technical report, 2015.

Olayinka A Onasoga, Joel A Afolayan, and Bukola D Oladimeij. Factors influencing utilization of antenatal care services among pregnant women in ife central lga, osun state, nigeria. *Advances in Applied Science Research*, 3(3):1309–1315, 2012.

Philip Oreopoulos, Marianne E Page, and Ann Huff Stevens. The intergenerational effects of compulsory schooling. *Journal of Labor Economics*, 24(4):729–760, 2006.

Philip Oreopoulos, Mark Stabile, Randy Walld, and Leslie L Roos. Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins. *Journal of Human Resources*, 43(1):88–138, 2008.

Emily Oster. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 2014.

Emily Oster and Rebecca Thornton. Menstruation, sanitary products, and school attendance: Evidence from a randomized evaluation. *American Economic Journal: Applied Economics*, 3(1):91–100, 2011.

Mary D Overpeck, Ruth A Brenner, Ann C Trumble, Lara B Trifiletti, and Heinz W Berendes. Risk factors for infant homicide in the united states. *New England Journal of Medicine*, 339(17):1211–1216, 1998.

C. Padez. Age at menarche of schoolgirls in maputo, mozambique. *Annals of Human Biology*, 30(4):487–495, 2003. doi: 10.1080/0301446031000111401.

Dillee Prasad Paudel, Baburao Nilgar, and Manisha Bhandankar. Determinants of postnatal maternity care service utilization in rural belgaum of karnataka, india: A community based cross-sectional study. *International Journal of Medicine and Public Health*, 4(1), 2014.

Nicola Persico, Andrew Postlewaite, and Dan Silverman. The effect of adolescent experience on labor market outcomes: The case of height. *Journal of Political Economy*, 112(5):1019–1053, 2004.

Rajesh Kumar Rai, Wafaie W Fawzi, Anamitra Barik, Abhijit Chowdhury, et al. The burden of iron-deficiency anaemia among women in india: how have iron and folic acid interventions fared? *WHO South-East Asia journal of public health*, 7(1):18, 2018.

Anita Raj, Niranjan Saggurti, Donta Balaiah, and Jay G Silverman. Prevalence of child marriage and its effect on fertility and fertility-control outcomes of young women in india: a cross-sectional, observational study. *The Lancet*, 373 (9678):1883–1889, 2009.

Nitya Rao. Total literacy campaigns: a field report. *Economic and Political Weekly*, pages 914–918, 1993.

Tomi-Ann Roberts, Jamie L Goldenberg, Cathleen Power, and Tom Pyszczynski. feminine protection: The effects of menstruation on attitudes towards women. *Psychology of Women Quarterly*, 26(2):131–139, 2002.

Joseph P Romano and Michael Wolf. Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469):94–108, 2005a.

Joseph P Romano and Michael Wolf. Stepwise multiple testing as formalized data snooping. *Econometrica*, 73(4):1237–1282, 2005b.

Joseph P Romano and Michael Wolf. Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics & Probability Letters*, 113:38–40, 2016.

Heather Royer. Separated at girth: Us twin estimates of the effects of birth weight. *American Economic Journal: Applied Economics*, 1(1):49–85, 2009.

Deborah Rubin, Cynthia P Green, and Altrena Mukuria. Addressing early marriage in uganda. 2009.

E Saar, C Shalev, I Dalal, and UA Sod-Moriah. Age at menarche: the influence of environmental conditions. *International Journal of Biometeorology*, 32(1):33–35, 1988.

Joseph J Sabia and Daniel I Rees. The effect of sexual abstinence on females' educational attainment. *Demography*, 46(4):695–715, 2009.

Joseph J Sabia and Daniel I Rees. Boys will be boys: are there gender differences in the effect of sexual abstinence on schooling? *Health Economics*, 20(3):287–305, 2011.

Bruce Sacerdote. The nature and nurture of economic outcomes. *American Economic Review*, 92(2):344–348, 2002.

Bruce Sacerdote. How large are the effects from changes in family environment? a study of korean american adoptees. *The Quarterly Journal of Economics*, 122 (1):119–157, 2007.

Shanika Samarakoon and Rasyad A Parinduri. Does education empower women? evidence from indonesia. *World Development*, 66:428–442, 2015.

Norbert Schady, Jere Behrman, Maria Caridad Araujo, Rodrigo Azuero, Raquel Bernal, David Bravo, Florencia Lopez-Boo, Karen Macours, Daniela Marshall, Christina Paxson, et al. Wealth gradients in early childhood cognitive development in five latin american countries. *Journal of Human Resources*, 50(2), 2015.

T Paul Schultz. Women's changing participation in the labor force: a world perspective. *Economic Development and Cultural Change*, 38(3):457–488, 1990.

Rachel K. Sebudde, Quentin T. Wodon, and Joseph Mawejje. Uganda economic update and accelerating uganda's development : ending child marriage, educating girls. *World Bank*, 2017.

Sheetal Sekhri and Sisir Debnath. Intergenerational consequences of early age marriages of girls: Effect on childrens human capital. *The Journal of Development Studies*, 50(12):1670–1686, 2014.

Enrique Sentana and Gabriele Fiorentini. Identification, estimation and testing of conditionally heteroskedastic factor models. *Journal of Econometrics*, 102(2): 143–164, 2001.

Manisha Shah and Bryce Millett Steinberg. Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy*, 125(2):527–561, 2017.

Mary A Silles. The causal effect of education on health: Evidence from the united kingdom. *Economics of Education review*, 28(1):122–128, 2009.

Mary A Silles. The intergenerational transmission of education: New evidence from adoptions in the usa. *Economica*, 84(336):748–778, 2017.

Bibha Simkhada, Edwin R van Teijlingen, Maureen Porter, and Padam Simkhada. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. *Journal of Advanced Nursing*, 61 (3):244–260, 2008.

Abhijeet Singh. Learning more with every year: School year productivity and international learning divergence. 2017.

Kitae Sohn. The height premium in indonesia. *Economics & Human Biology*, 16: 1–15, 2015.

Marni Sommer. Where the education system and women's bodies collide: The social and health impact of girls' experiences of menstruation and schooling in tanzania. *Journal of Adolescence*, 33(4):521–529, 2010.

Henrik Toft Sørensen, Svend Sabroe, Kenneth J Rothman, Matthew Gillman, Flemming Hald Steffensen, Peer Fischer, and Thorkild IA Serensen. Birth weight and length as predictors for adult height. *American Journal of Epidemiology*, 149(8):726–729, 1999.

Kaspar Sørensen, Anders Juul, Kaare Christensen, Axel Skytthe, Thomas Scheike, and Tina Kold Jensen. Birth size and age at menarche: a twin perspective. *Human Reproduction*, 28(10):2865–2871, 2013.

Dean Spears. Height and cognitive achievement among indian children. *Economics & Human Biology*, 10(2):210–219, 2012.

Douglas Staiger and JH Stock. Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586, 1997.

Philip J Steer. Maternal hemoglobin concentration and birth weight–. *The American Journal of Clinical Nutrition*, 71(5):1285S–1287S, 2000.

RJ Stoltzfus, LC Mullany, and RE Black. Iron deficiency anemia, comparative quantification of health risks. *Geneva: WHO*, 2004.

Richard S Strauss. Effects of the intrauterine environment on childhood growth. *British Medical Bulletin*, 53(1):81–95, 1997.

SV Subramanian, Emre Özaltin, and Jocelyn E Finlay. Height of nations: a socioeconomic analysis of cohort differences and patterns among women in 54 low-to middle-income countries. *PLoS One*, 6(4):e18962, 2011.

Naveen Sunder. Marriage age, social status and intergenerational effects in uganda. Technical report, Working Paper, 2018a.

Naveen Sunder. Schooling & health : Causal estimates from a school construction programme in india. *Working Paper*, 2018b.

James M Tanner. A concise history of growth studies from buffon to boas. In *Human Growth*, pages 515–593. Springer, 1979.

Georgia Taylor and Paola Pereznieto. Review of evaluation approaches and methods used by interventions on women and girls' economic empowerment. *Overseas Development Institute,(ODI), London*, 2014.

Petra E Todd and Kenneth I Wolpin. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113 (485):F3–F33, 2003.

Petra E Todd and Kenneth I Wolpin. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, 1(1):91–136, 2007.

Petia Topalova. Trade liberalization, poverty and inequality: Evidence from indian districts. In *Globalization and Poverty*, pages 291–336. University of Chicago Press, 2007.

UNICEF. Child marriage: Latest trends and future prospects. *New York: UNICEF*, 2018.

Cesar G Victora, Linda Adair, Caroline Fall, Pedro C Hallal, Reynaldo Martorell, Linda Richter, Harshpal Singh Sachdev, Maternal, Child Undernutrition Study Group, et al. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet*, 371(9609):340–357, 2008.

Tom S Vogl. Height, skills, and labor market outcomes in mexico. *Journal of Development Economics*, 107:84–96, 2014.

Sebastian Vollmer, Christian Bommer, Aditi Krishna, Kenneth Harttgen, and SV Subramanian. The association of parental education with childhood undernutrition in low-and middle-income countries: comparing the role of paternal and maternal education. *International Journal of Epidemiology*, 46(1):312–323, 2016.

Huan Wang, James Chu, Prashant Loyalka, Xin Tao, Yaojiang Shi, Qinghe Qu, Chu Yang, and Scott Rozelle. Can school counseling reduce school dropout in developing countries? *Rural Education Action Program Working Paper No*, 275, 2014.

Quentin Wodon, Chata Male, Adenike Onagoruwa, Aboudrahyme Savadogo, and Ali Yedan. Child marriage, early childbearing, low educational attainment for girls, and their impacts in uganda, 2017.

WorldBank. World bank dataverse. *World Bank Dataverse*, 2016.

Ji Yan. Maternal pre-pregnancy bmi, gestational weight gain, and infant birth weight: A within-family analysis in the united states. *Economics & Human Biology*, 18:1–12, 2015.

Joanne Yoong. The impact of economic resource transfers to women versus men: a systematic review. 2012.

Kathryn M Yount, Kristin E VanderEnde, Sylvie Dodell, and Yuk Fai Cheong. Measurement of womens agency in egypt: A national validation study. *Social Indicators Research*, 128(3):1171–1192, 2016.

Kathryn M Yount, AliceAnn Crandall, and Yuk Fai Cheong. Womens age at first marriage and long-term economic empowerment in egypt. *World Development*, 102:124–134, 2018.

Desalegn T Zegeye, Berihun Megabiaw, and Abay Mulu. Age at menarche and the menstrual pattern of secondary school adolescents in northwest ethiopia. *BMC Women's Health*, 9(1):29, 2009.